The evolutionary ecology of antibiotic resistance and antibiotic resistance plasmids in *Escherichia coli*

Frances Anna Medaney

Royal Holloway University of London

Submitted for the degree of Doctor of Philosophy

Declaration of Authorship

I Frances Anna Medaney hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, this is always clearly stated.

Signed:

Date:

Abstract

Antibiotic resistance is '*as big a risk as terrorism'* according to H.M. Government Chief Medical Officer in March 2013. New insights into the ecology and evolution of resistance are, therefore, a research priority if we are to avert this crisis. Antibiotic resistance genes are often carried on plasmids, mobile genetic elements that allow rapid dissemination of resistance through bacterial populations. However, little is known about the ecology of plasmids and associated antibiotic resistance in natural bacterial communities. In addition, detoxifying antibiotic resistance is often cited as a social trait in bacteria. This is of particular interest because many virulence-associated traits, including siderophore production and toxin secretion, are cooperative. This thesis examines the evolutionary ecology of both antibiotic resistance and resistance plasmids in *E. coli* in four related studies: i) an examination cooperative β-lactam resistance using laboratory competition experiments, which demonstrated that cooperative resistance occurs only under very specific conditions: the presence of 'persister' susceptible colonies, which are not resistant, but can tolerate high concentrations of antibiotic in the presence of resistant cells; ii) sequencing of non-resistance plasmids, which demonstrated the wide variety of plasmids present in one bacterial community, and highlighted our limited understanding of plasmid gene function; iii) a population study of *E. coli* and its plasmids which found surprisingly low levels of antibiotic resistance, and interesting patterns of plasmid-host diversity and population structuring; and iv) laboratory plasmid maintenance experiments which found a complex, frequency-dependent pattern governing plasmid persistence.

These results bring into question current assumptions about resistance as a cooperative trait, and suggest that frequency-dependence could be key to explaining plasmid maintenance dynamics. They also provide insights into natural plasmid populations through sequencing and field study, which contribute to our understanding of the spread of antibiotic resistance genes.

Acknowledgements

Firstly, thanks to my supervisors, Ben Raymond and Richard Ellis, and to Royal Holloway University of London and Animal Health and Veterinary Laboratories Agency for funding this work. Also thanks to all my colleagues and fellow students at both institutions, in particular Dr Liqin Zhou and (soon to be Dr) Andrew Matthews.

Thanks to my family, and to TW, and finally thanks to my friends, if I have any left…

"The road goes ever on and on…"

JRR Tolkien, as quoted by JF

"A contribution to knowledge"

For SM

"Anyone who sees me go anywhere near a university again,
you've got my permission to shoot me."

TW, paraphrasing Sir Steve Redgrave

Table of Contents

Table of Figures

CHAPTER 6

Table of Tables

Chapter 1: General Introduction

1.1 Introductory statement

Over twenty years ago Harold Neu published '*The Crisis in Antibiotic Resistance*' (Neu 1992), which described the increasing prevalence of resistance across many bacterial genera and cited the overuse of antibiotics in hospitals, in the community and in animal husbandry as key contributory factors. Over the subsequent decades 'superbugs' such as methicillin-resistant *Staphylococcus aureus* (MRSA) have continued to pose a huge challenge to medical professionals (Mulvey & Simor 2009). The antibiotic resistance issue has even begun to permeate the national psyche, as demonstrated by recent statements from government (The Department of Health & The Prime Minister's Office, July 2014), health officials (Davies 2013) and the general public, who have voted to make 'Antibiotics' the winning challenge for the Longitude Prize 2014 (see http://www.longitudeprize.org).

Bacteria have so clearly demonstrated their capacity to evolve resistance to new drugs that alternative strategies for resistance management are desperately needed: a new drug will only remain effective if resistance can be controlled. Antibiotic resistance genes are often carried on plasmids, mobile genetic elements that allow rapid dissemination of resistance through bacterial populations. This thesis explores a novel strategy for exploiting plasmids and bacterial cooperation to limit the spread of resistance, and offers insights into the evolutionary ecology of plasmids and antibiotic resistance in natural bacterial communities. The introduction provides some background about antibiotic resistance and resistance plasmids, focusing on cooperative β-lactamases and leading to the 'cheat' plasmid hypothesis and its potential for resistance management. As well as exploring strategies for limiting antibiotic resistance, this thesis addresses more general questions of plasmid ecology and persistence which are outlined below, alongside an overview of the study system and the specific aims of this research.

1.2 Antibiotic resistance and resistance plasmids

1.2.1 Mechanisms of resistance

Antibiotics have many varied modes of action, so it follows that resistance mechanisms too are varied; resistance can be conferred by a single point mutation in the target gene, or it may require the acquisition of a whole new set of genes (Mazel 2001; Hayes & Wolf 1990). Figure 1.1 illustrates five key resistance strategies: i) reduced drug import due to alterations in the cell membrane or in active transporters such as *omp* proteins (Hayes & Wolf 1990); ii) target site modification, such as rifampicin resistance due to mutations in the RNA polymerase β-subunit (Spratt 1994); iii) drug modification by cleavage of the drug (β-lactams) or by enzymatic alterations such as acetylation or phosphorylation (kanamycin) (Walsh 2000); iv) antibiotic sequestration, such as bleomycin binding proteins (Dumas et al. 1994); and finally v) drug efflux, the active removal of antibiotics from the cell by general or specific pumps, which is often associated with multi-drug resistance (Webber & Piddock 2003).

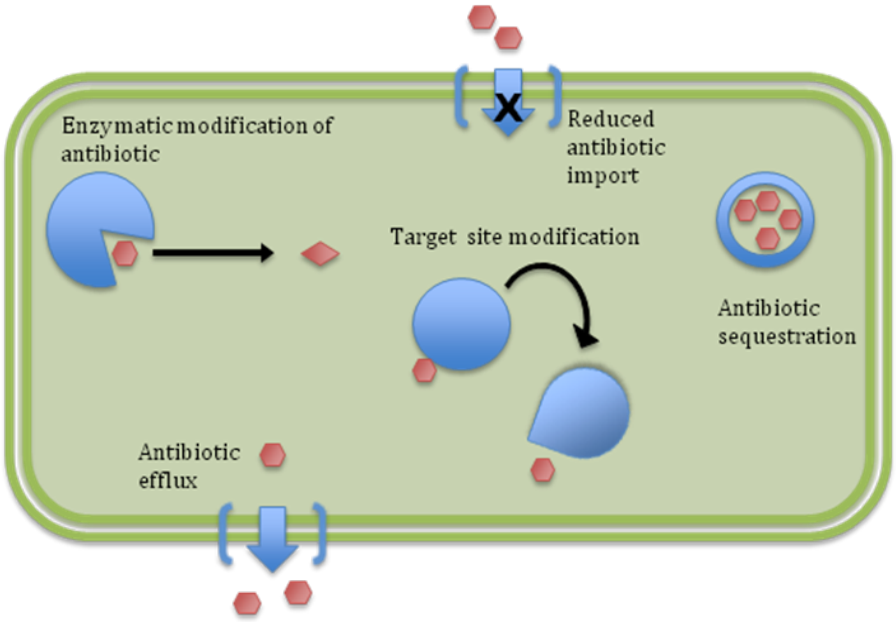

Figure 1.1: Mechanisms of antibiotic resistance.

1.2.2 β-lactam action and resistance

Antibiotics containing a β-lactam ring, penicillins and cephalosporins, target the bacterial cell wall. Much of the mechanism behind this process remains unclear despite over seven decades of research. However, it is known that the drugs bind to penicillin-binding proteins (PBPs, which are many and varied) preventing peptidoglycan synthesis and stimulating the release of autolysins (Blumberg & Strominger 1974; Waxman & Strominger 1983; Lewis 2007). This produces a range of outcomes depending on the PBP affected, the strain and the growth conditions; cell lysis, production of distorted round-form cells, production of filaments and combinations of the three are possible (Gould & MacKenzie 1997). Lysis is clearly a critical factor for cell death. Penicillins are bactericidal to growing cells with the rate of killing proportional to growth rate (Hobby et al. 1942; Rolinson et al. 1977), but they are largely ineffective on non-growing cells (with limited exceptions, see Tuomanen et al. 1986).

The first β-lactam resistance enzyme was identified before the release of penicillin for medical use (Abraham & Chain 1940; Bradford 2001). These enzymes, the β-lactamases, confer resistance by facilitating cleavage of the β-lactam ring hydrolysing penicillins to penicilloic acids (Figure 1.2). This hydrolysis inhibits antibiotic binding to PBPs, rendering the drug inactive (Walsh 2000; Davies 1994). Extended-spectrum β-lactamases (ESBLs), which confer cephalosporin as well as penicillin resistance, are now widespread across Europe, and pose an increasing risk to both human and animal health (Livermore et al. 2008; Canton et al. 2008; Bradford 2001).

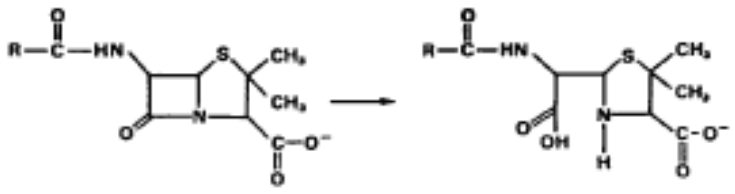

Figure 1.2: Hydrolysis of penicillin to penicilloic acid, catalysed by β-lactamase enzymes. Adapted from Sykes & Matthew (1976).

1.2.3 Resistance plasmids and the spread of ESBLs

Horizontal gene transfer, the lateral exchange of genes between individuals, both within and between species, has been identified as one of the major factors in the spread of antibiotic resistance genes through bacterial populations (de la Cruz & Davies 2000). In particular, conjugation, plasmid transfer via cell to cell contact, is critical for the spread of resistance genes (Ochman et al. 2000). Plasmids are major carriers of antibiotic resistance genes: they facilitate gene transfer across bacterial genera, and allow multiple resistance genes to accrue on one transmissible element (Taylor et al. 2004; Davies 1994). They are also diverse, encompassing large and small, high- and low- copy number and many incompatibility (Inc) groups: plasmids with closely related replication control genes cannot be maintained in the same host cell and are termed incompatible (Taylor et al. 2004).

The spread of ESBLs, in particular CTX-M β-lactamases, has been largely plasmid mediated (Canton et al. 2008). Plasmid carried CTX-M genes are known to be derived from chromosomal β-lactamases of *Kluyvera sp*. via insertion sequence mobilization (Livermore et al. 2006). The wide variety of plasmids, many of which carry additional resistance genes, may indicate multiple mobilization events: one IncF plasmid, pEK499, was found to carry a cluster of ten resistance genes, including CTX-M-15 (Woodford et al. 2009). Worryingly, these CTX-M carrying plasmids have also been linked to virulent *E. coli* genotypes (Cantón & Coque 2006; Rossolini et al. 2008). The emergence of CTX-M enzymes in *Enterbacteriaceae* since the 1980s has seen a rapid increase in the prevalence of resistance and changes in plasmid epidemiology in both humans and livestock (Dierikx et al. 2010; Bou et al. 2002; Diestra et al. 2009). Moreover, there is evidence for ESBL plasmid transmission from animals to humans via the food chain, highlighting the importance of resistance management in livestock as well as humans (Leverstein-van Hall et al. 2011).

1.3 Microbial cooperation and resistance

1.3.1 Microbial cooperation

West et al. (2006) define cooperation as "A behaviour that benefits another individual (the recipient) and which is maintained (at least partially) because of its beneficial effect on the recipient". From an evolutionary perspective, these behaviours are hard to justify, as they appear to be fundamentally at odds with Darwin's concept of 'survival of the fittest'. However, this paradigm has been largely explained by Hamilton's rule, which states that altruistic (cooperative) acts are favoured when the cost to the altruist (c) is low and the benefit to the recipient (b) and relatedness (r) of the two is high:

$$rb - c > 0$$

(Frank 1998; Hamilton 1964; West et al. 2006). By directing altruistic acts towards relatives, individuals may still gain indirect fitness benefits, by improving the reproductive success of individuals who are likely to carry the same genes (West et al. 2007). Indirect fitness is dependent on altruism being directed towards individuals with shared genes by the process of kin selection. Two mechanisms govern this: limited dispersal and kin discrimination (Smith 1964). Kin discrimination requires recognition of relatives so that altruistic acts can be directed towards them, whereas limited dispersal involves indiscriminate altruism towards close neighbours, assuming that they are also likely to be close relatives (West et al. 2007). Limited dispersal is particularly relevant in microbial communities, where cells reproduce clonally so that close neighbours are almost certainly close relatives.

In bacteria, cooperative traits often come in the form of 'public goods' released into the environment and available to all. These traits are open to exploitation by selfish individuals, social 'cheats', who benefit from public goods but do not contribute to their production (and therefore do not pay any fitness costs). Ultimately this can lead to a collapse of cooperation and the 'tragedy of the commons' where each individual acts selfishly to the detriment of the population

as a whole (Hardin 1968). Many virulence-associated traits such as quorum sensing, siderophore production and toxin secretion are cooperative (Crespi 2001; Raymond & Bonsall 2013; Griffin et al. 2004), and it has been suggested that these traits could be a target for antimicrobial therapy (André & Godelle 2005) .

1.3.2 Cooperative resistance & 'cheat' plasmids

β-lactamase enzymes can be considered to be public goods, as they benefit the whole bacterial population by detoxifying the environment. This cooperative resistance is open to exploitation by social cheats; cells that benefit from the enzyme produced by cooperators, but which do not themselves produce the public good. It has been shown that antibiotic sensitive bacteria can survive high concentrations of antibiotic in the presence of resistant strains (Dugatkin et al. 2005; Perlin et al. 2009; Clark et al. 2009).

Antibiotic resistance genes are often carried on plasmids. Horizontal transfer of plasmids to plasmid-free cells would ensure that all cells carry the resistance gene and share the cost of public good production, thereby preventing the evolution of social cheats. However, plasmid spread is affected by plasmid incompatibility: cells cannot maintain two plasmids of the same incompatibility group (Taylor et al. 2004). The 'cheat' plasmid hypothesis predicts the existence of plasmids with similar cost to the resistance plasmid in the same incompatibility group but lacking the resistance gene. Carriage of a 'cheat' plasmid would prevent cells from acquiring the resistance plasmid, thus maintaining susceptible cheats in the population despite horizontal gene transfer (HGT). It has been proposed that 'Trojan horse' susceptible cheats could invade and displace resistant populations, facilitating clearance of an infection with antibiotics (Brown et al. 2009). 'Cheat' plasmids could block the spread of resistance genes by horizontal transfer, and thus invade resistant populations.

1.4 Plasmid populations and persistence

1.4.1 Plasmid populations

It is often stated that plasmids act as selfish genetic elements, existing purely for their own replication and proliferation (Werren 2011). The existence of post-segregational killing systems (toxin-antitoxin systems), where a plasmid-encoded toxin kills segregant daughter cells that have lost the corresponding antitoxin, is often cited to support the idea of selfish plasmids. However, it has also been proposed that these toxin-antitoxin systems may have evolved as a mechanism for plasmid-plasmid competition (Van Melderen & Saavedra De Bast 2009; Kobayashi 2004). These conflicting theories highlight our limited understanding of plasmid ecology and the selective forces driving plasmid evolution: plasmids exist not as individual entities, but as parts of communities interacting with each other and with bacterial hosts. Few studies have examined plasmid communities as a whole, and those which have often focus on plasmids associated with virulence and disease, in particular antibiotic resistance (Carattoli 2011; Johnson & Nolan 2009). This not only presents a biased view of plasmid populations and their dynamics but also fails to explain the existence of 'cryptic' plasmids, which appear to have no functional genes at all. A broader understanding of plasmid communities could inform strategies for virulence management, such as 'cheat' plasmids, and the 'Trojan horse' approach proposed by Brown et al. (2009), as discussed in Section 1.3.2 above.

1.4.2 Plasmid persistence

Plasmid carriage is widely assumed to be costly to the host. This can be a metabolic cost due to increased DNA replication and maintenance (especially for large or high copy number plasmids) or due to expression of plasmid genes. It can also be a phenotypic cost such as the expression of conjugation apparatus, which prevents replication and can be a target for bacteriophage (Thomas 2004; Eberhard 1990). Moreover, costly plasmids may be lost through segregation; uneven distribution during cell division. Given the cost of plasmid carriage and the rate of segregation, it is easy to imagine that plasmids would be lost from the population via segregation and competition with fitter plasmid-free cells in the

absence of selection in their favour. Conjugation may mitigate these factors by transferring plasmids to new hosts, and theory predicts that plasmids can be maintained parasitically (without conferring a benefit to the host) if these three key parameters can be balanced appropriately, i.e. high conjugation, low cost and low segregation rate, summarised by the equation (Stewart & Levin 1977; Levin & Stewart 1980):

$$\text{Transfer Rate} > \text{Cost of Carriage} + \text{Segregation Rate}$$

The rise of multi-resistance plasmids, and the fact that many virulence associated genes are plasmid associated, heightens the importance of understanding plasmid maintenance if we are to limit their spread. The debate as to whether virulence or antibiotic resistance plasmids can survive in the absence of selection is also be key in the development of containment and management strategies in healthcare.

1.5 Aims of this thesis

1.5.1 Introduction to the study system

This thesis uses *E. coli* and *E. coli* plasmids, isolated from cattle, to address evolutionary ecology questions for both antibiotic resistance and resistance plasmids. Antibiotic resistance is as pressing an issue for veterinarians as for clinicians; moreover, there is evidence of plasmid and bacterial transmission from animals to humans, which highlights the importance of resistance management in livestock (Leverstein-van Hall et al. 2011; Aslam et al. 2003; Liebana et al. 2004). Chapters 2, 3 and 5 use the resistance plasmid pCT, isolated by Liebana et al. (2006) from the site of the first veterinary ESBL CTX-M *E. coli* outbreak in the UK. pCT is a large, ~93Kbp plasmid, carrying the extended spectrum β-lactamase, CTX-M-14. Extended-spectrum β-lactamases (ESBLs), which confer cephalosporin as well as penicillin resistance, are now widespread, with CTX-M-type ESBLs becoming the dominant type in Europe over the last decade (see Section 1.2.3 above), however their spread in livestock has not been as rapid as that seen in humans (Liebana et al. 2006). In addition to pCT, four

other plasmids from the same study site are investigated, alongside a field study that collected new *E. coli* and plasmids from cattle.

1.5.2 Aims and thesis outline

The original aims of this thesis were as follows: i) to examine the conditions under which β-lactam resistance is cooperative; ii) to identify and sequence 'cheat' plasmids from natural zoonotic bacterial populations; iii) to investigate the manipulation of cooperative resistance population dynamics using 'cheat' plasmids. Chapter 2 provides details of mutant strains and plasmids produced for this thesis, as well as generic methodologies. The first aim is addressed in Chapter 3, which demonstrates the very particular conditions required for cooperative resistance. The second aim, to identify and sequence potential 'cheat' plasmids, is addressed in Chapters 4 and 5. Chapter 5 also includes a non-selective field study of a zoonotic plasmid population. The work in Chapters 2 to 5 led to a further aim; iv) to investigate mechanisms of plasmid persistence, in particular the role of conjugation in the maintenance of large naturally occurring plasmids. This is addressed in Chapter 6.

Plasmids are major carriers of antibiotic resistance genes. A better understanding of their persistence, population dynamics and spread is critical if we are to limit the resistance crisis. Furthermore, bacteria are always going to evolve resistance to new drugs, so it is therefore vital to explore alternative ways of tackling infections, such as exploiting their own population dynamics (Brown et al. 2009), or targeting plasmids themselves (Williams & Hergenrother 2008). In addition to the resistance issue, plasmids are interesting in their own right: most studies have focused on plasmids carrying resistance or other virulence factors, and although this is important, it gives a skewed view of plasmid populations. To truly understand plasmid persistence, non-selective strategies must be employed.

References

Abraham, E.P. & Chain, E., 1940. An enzyme from bacteria able to destroy penicillin. *Nature*, 146(3713), pp.837–837.

André, J. & Godelle, B., 2005. Multicellular organization in bacteria as a target for drug therapy. *Ecology Letters*, 8(8), pp.800–810.

Anon, *The Longitude Prize 2014*, Available at: http://www.longitudeprize.org [Accessed July 5, 2014].

Aslam, M. et al., 2003. Origin of contamination and genetic diversity of *Escherichia coli* in beef cattle. *Applied and Environmental Microbiology*, 69(5), pp.2794–2799.

Blumberg, P. & Strominger, J., 1974. Interaction of penicillin with the bacterial cell: penicillin-binding proteins and penicillin-sensitive enzymes. *Bacteriological Reviews*, 38(3), pp.291–335.

Bou, G. et al., 2002. Identification and broad dissemination of the CTX-M-14 β-lactamase in different *Escherichia coli* strains in the northwest area of Spain. *Journal of clinical microbiology*, 40(11), pp.4030–4036.

Bradford, P., 2001. Extended-spectrum β-lactamases in the 21st century: Characterization, epidemiology, and detection of this important resistance threat. *Clinical Microbiology Reviews*, 14(4), pp.933–951.

Brown, S. et al., 2009. Social evolution in micro-organisms and a Trojan horse approach to medical intervention strategies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364, pp.3157–3168.

Canton, R. et al., 2008. Prevalence and spread of extended-spectrum β-lactamase-producing *Enterobacteriaceae* in Europe. *Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, 14(s1), pp.144–153.

Cantón, R. & Coque, T.M., 2006. The CTX-M β-lactamase pandemic. *Current Opinion in Microbiology*, 9(5), pp.466–475.

Carattoli, A., 2011. Plasmids in Gram negatives: Molecular typing of resistance plasmids. *International Journal of Medical Microbiology*, 301(8), pp.654–658.

Clark, D., Alton, T. & Bajorek, A., 2009. Evolution of altruists and cheaters in near-isogenic populations of *Escherichia coli*. *Frontiers in Bioscience (Landmark edition)*, 14, pp.4815–4824.

Crespi, B., 2001. The evolution of social behavior in microorganisms. *Trends in Ecology & Evolution*, 16(4), pp.178–183.

Davies, J., 1994. Inactivation of antibiotics and the dissemination of resistance genes. *Science*, 264(5157), pp.375–382.

Davies, S.C., 2013. *Annual report of the Chief Medical Officer: Infections and the rise of antimicrobial resistance* D. Walker & T. Fowler, eds., The Department of Health.

de la Cruz, F. & Davies, J., 2000. Horizontal gene transfer and the origin of species: lessons from bacteria. *Trends in Microbiology*, 8(3), pp.128–133.

Dierikx, C. et al., 2010. Increased detection of extended spectrum β-lactamase producing *Salmonella enterica* and *Escherichia coli* isolates from poultry. *Veterinary Microbiology*, 145(3-4), pp.273–278.

Diestra, K., Juan, C. & Curiao, T., 2009. Characterization of plasmids encoding blaESBL and surrounding genes in Spanish clinical isolates of *Escherichia coli* and *Klebsiella pneumoniae*. *Journal of Antimicrobial Chemotherapy*, 63, pp.60–66.

Dugatkin, L. et al., 2005. Group-beneficial traits, frequency-dependent selection and genotypic diversity: an antibiotic resistance paradigm. *Proceedings of the Royal Society B: Biological Sciences*, 272, pp.79–83.

Dumas, P. et al., 1994. Crystal structure and site-directed mutagenesis of a bleomycin resistance protein and their significance for drug sequestering. *The EMBO Journal*, 13(11), pp.2483–2492.

Eberhard, W., 1990. Evolution in bacterial plasmids and levels of selection. *The Quarterly review of biology*, 65(1), pp.3–22.

Frank, S., 1998. *Foundations of social evolution* 1st ed. J. R. Krebs & T. Clutton-Brock, eds., Princeton, USA: Princeton University Press.

Gould, I.M. & MacKenzie, F.M., 1997. The response of *Enterobacteriaceae* to β-lactam antibiotics 'round forms, filaments and the root of all evil'. *Journal of Antimicrobial Chemotherapy*, 40, pp.495-499.

Griffin, A.S., West, S.A. & Buckling, A., 2004. Cooperation and competition in pathogenic bacteria. *Nature*, 430(7003), pp.1024–1027.

Hamilton, W., 1964. The genetical evolution of social behaviour. II. *Journal of theoretical biology*, 7(1), pp.17–52.

Hardin, G., 1968. The tragedy of the commons. *Science*, 162, pp.1243–1248.

Hayes, J. & Wolf, C., 1990. Molecular mechanisms of drug resistance. *Biochemical journal*, 272(2), pp.281–295.

Hobby, G.L., Meyer, K. & Chaffee, E., 1942. Observations on the mechanism of action of penicillin. *Experimental Biology and Medicine*, 50(2), pp.281–285.

Johnson, T.J. & Nolan, L.K., 2009. Pathogenomics of the virulence plasmids of *Escherichia coli*. *Microbiology and molecular biology reviews: MMBR*, 73(4), pp.750–774.

Kobayashi, I., 2004. Genetic addiction: a principle of gene symbiosis in a genome. In B. E. Funnell & G. J. Phillips, eds. *Plasmid Biology*. Washington D.C.: American Society for Microbiology (ASM), pp. 105–144.

Leverstein-van Hall, M.A. et al., 2011. Dutch patients, retail chicken meat and poultry share the same ESBL genes, plasmids and strains. *Clinical Microbiology and Infection*, 17(6), pp.873–880.

Levin, B.R. & Stewart, F.M., 1980. The population biology of bacterial plasmids: a priori conditions for the existence of mobilizable nonconjugative factors. *Genetics*, 94(2), pp.425–443.

Lewis, K., 2007. Persister cells, dormancy and infectious disease. *Nature reviews Microbiology*, 5(1), pp.48–56.

Liebana, E. et al., 2004. Characterization of β-lactamases responsible for resistance to extended-spectrum cephalosporins in *Escherichia coli* and *Salmonella enterica* strains from food-producing animals in the United Kingdom. *Microbial Drug Resistance*, 10(1), pp.1–9.

Liebana, E. et al., 2006. Longitudinal farm study of extended-spectrum β-lactamase-mediated resistance. *Journal of clinical microbiology*, 44(5), pp.1630–1634.

Livermore, D.M. et al., 2006. CTX-M: changing the face of ESBLs in Europe. *Journal of Antimicrobial Chemotherapy*, 59(2), pp.165–174.

Livermore, D.M. et al., 2008. Non-susceptibility trends among *Enterobacteriaceae* from bacteraemias in the UK and Ireland, 2001-06. *Journal of Antimicrobial Chemotherapy*, 62 (Suppl. 2), pp.ii41–54.

Mazel, D., 2001. Bacterial genetics and antibiotic resistance dissemination. In D. Hughes & D. I. Anderson, eds. *Antibiotic Development and Resistance*. London: Taylor and Francis, pp. 91–106.

Mulvey, M.R. & Simor, A.E., 2009. Antimicrobial resistance in hospitals: how concerned should we be? *CMAJ: Canadian Medical Association journal*, 180(4), pp.408–415.

Neu, H.C., 1992. The crisis in antibiotic resistance. *Science*, 257(5073), pp.1064–1073.

Ochman, H., Lawrence, J.G. & Groisman, E.A., 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature*, 405(6784), pp.299–304.

Perlin, M.H. et al., 2009. Protection of *Salmonella* by ampicillin-resistant *Escherichia coli* in the presence of otherwise lethal drug concentrations. *Proceedings of the Royal Society B: Biological Sciences*, 276(1674), pp.3759–3768.

Raymond, B. & Bonsall, M.B., 2013. Cooperation and the evolutionary ecology of

bacterial virulence: the *Bacillus cereus* group as a novel study system. *BioEssays*, 35(8), pp.706–716.

Rolinson, G., Macdonald, A. & Wilson, D., 1977. Bactericidal action of β-lactam antibiotics on *Escherichia coli* with particular reference to ampicillin and amoxycillin. *Journal of Antimicrobial Chemotherapy*, 3(6), pp.541–553.

Rossolini, G.M., D'Andrea, M.M. & Mugnaioli, C., 2008. The spread of CTX-M-type extended-spectrum β-lactamases. *Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, 14(Suppl. 1), pp.33–41.

Smith, J., 1964. Group selection and kin selection. *Nature*, 201, pp.1145–1147.

Spratt, B.G., 1994. Resistance to antibiotics mediated by target alterations. *Science*, 264(5157), pp.388–393.

Stewart, F.M. & Levin, B.R., 1977. The population biology of bacterial plasmids: a priori conditions for the existence of conjugationally transmitted factors. *Genetics*, 87(2), pp.209–228.

Sykes, R. & Matthew, M., 1976. The β-lactamases of Gram-negative bacteria and their role in resistance to β-lactam antibiotics. *Journal of Antimicrobial Chemotherapy*, 2(2), pp.115–157.

Taylor, D.E. et al., 2004. Antibiotic resistance plasmids. In B. E. Funnell & G. J. Phillips, eds. *Plasmid Biology*. Washington D.C. : American Society for Microbiology (ASM), pp. 473–491. A

The Department of Health & The Prime Minister's Office, July 2014. *Prime Minister warns of global threat of antibiotic resistance*.

Thomas, C.M., 2004. Evolution and population genetics of bacterial plasmids. In B. E. Funnell & G. J. Phillips, eds. *Plasmid Biology*. Washington: American Society for Microbiology (ASM), pp. 509–528.

Tuomanen, E. et al., 1986. The rate of killing of *Escherichia coli* by β-lactam antibiotics is strictly proportional to the rate of bacterial growth. *Microbiology*, 132(5), pp.1297–1304.

Van Melderen, L. & Saavedra De Bast, M., 2009. Bacterial toxin–antitoxin systems: more than selfish entities? *PLoS Genetics*, 5(3), p.e1000437.

Walsh, C., 2000. Molecular mechanisms that confer antibacterial drug resistance. *Nature*, 406(6797), pp.775–781.

Waxman, D.J. & Strominger, J., 1983. Penicillin-binding proteins and the mechanism of action of β-lactam antibiotics. *Annual Reviews Biochemistry*, 52, pp.825–869.

Webber, M.A. & Piddock, L.J.V., 2003. The importance of efflux pumps in bacterial

antibiotic resistance. *Journal of Antimicrobial Chemotherapy*, 51(1), pp.9–11.

Werren, J.H., 2011. Selfish genetic elements, genetic conflict, and evolutionary innovation. *Proceedings of the National Academy of Sciences of the United States of America*, 108(Suppl. 2), pp.10863–10870.

West, S.A. et al., 2006. Social evolution theory for microorganisms. *Nature reviews Microbiology*, 4(8), pp.597–607.

West, S.A. et al., 2007. The social lives of microbes. *Annual review of ecology, evolution, and systematics*, 38(1), pp.53–77.

Williams, J.J. & Hergenrother, P.J., 2008. Exposing plasmids as the Achilles' heel of drug-resistant bacteria. *Current Opinion in Chemical Biology*, 12, pp.389–399.

Woodford, N. et al., 2009. Complete nucleotide sequences of plasmids pEK204, pEK499, and pEK516, encoding CTX-M enzymes in three major *Escherichia coli* lineages from the United Kingdom, all belonging to the international O25:H4-ST131 clone. *Antimicrobial Agents and Chemotherapy*, 53(10), pp.4472–4482.

Chapter 2: Production of mutants & generic methodology

2.1 Introduction

The plasmid used throughout this thesis, pCT, is a large, ~93 kb plasmid, carrying the extended spectrum β-lactamase CTX-M-14, that was identified in *E. coli* strain C159/11 by Teale et al. (2005, as cited by Liebana et al., 2006), and sequenced by Cottell et al. (2011a, 2011b). pCT is an IncK plasmid and a member of the IncI complex, which includes incompatibility groups I1, I2, Iγ, B, O and K (Cottell 2011b, Praszkier et al. 1991). It carries two conjugation operons, *tra* and *pil*, and a shufflon recombinase, which are characteristic of IncI complex plasmids (Cottell 2011b) and make pCT highly transmissible. In order to examine the effects of conjugation on plasmid maintenance (Chapter 6) it was necessary to construct a conjugation-null pCT mutant. In addition to this, a *lac* operon knockout of the host strain, *Escherichia coli* MG1655 (Coli Genetic Stock Center #7740, Yale University, Connecticut, USA), was constructed for use in Chapters 3 and 6. This second mutant provided a near-isogenic background for competition experiments, and facilitated differentiation between resistant and susceptible strains using blue/white screening in the presence of X-Gal & IPTG (5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside & isopropyl-beta-D-thiogalactopyranoside).

Cottell et al. (2011a) identify two sets of conjugation genes on pCT; the *tra* and *pil* loci, which show very high identity (99%) with the IncI plasmid R64. R64 has 23 genes essential for conjugation, including several *tra* and *trb* genes with homologues in pCT (Cottell et al. 2011a; Komano et al. 2000). In R64 the *pil* locus is not thought to be essential for all conjugation, but rather is just for liquid mating, so *pil* genes were discounted for deletion, as they would not fully inhibit conjugation. The *trbA* gene, which encodes a putative regulation protein, at position 38848-40110 on pCT, was selected for deletion, as it is "indispensible for R64 transfer" (Komano et al. 2000).

The MG1655 *lac* operon deletion used primers taken from Datsenko & Wanner (2000). These primers disrupt the chromosomal *lacZYA* genes (nucleotides 360842-365662), producing a *Δlac* mutant with a white phenotype in the presence of X-Gal & IPTG.

In addition to discussing the production and validation of the two mutants, this chapter will give details of the generic methodology used throughout this work including culture conditions, strain storage and basic molecular techniques. Any deviations from these procedures will be made clear in subsequent chapters.

2.2 Methods

2.2.1 Generic methodology

Culture media

Liquid cultures were grown in 5ml Luria Bertani (LB) broth (Fisher Scientific UK Ltd., Loughborough, UK). Plates were LB broth with 2% (w/v) agar (Agar Bacteriological No. 1, Oxoid, Basingstoke, UK). Broth and plates were supplemented with antibiotics as required; pCT plasmid carrying strains were maintained on plates containing either ampicillin or cefotaxime. See Table 2.1 for standard antibiotic doses used throughout this work.

Antibiotic	**Supplier**	**Dose (µg ml^{-1})**
Ampicillin sodium salt	Sigma-Aldrich Company Ltd., Gillingham, UK	100
Cefotaxime sodium salt	Melford Laboratories Ltd., Ipswich, UK	8
Chloramphenicol	Sigma-Aldrich	20
Tetracycline hydrochloride	Sigma-Aldrich	12
Rifampicin	Sigma-Aldrich	100
Nalidixic acid	Sigma-Aldrich	15

Table 2.1: Antibiotics used including standard doses and suppliers.

Overnight culture

Overnight (o/n) cultures were initiated from single colonies on fresh LB agar plates. Colonies were inoculated into 5ml LB broth and incubated, shaken at 200 rpm, at 37°C for approximately 16 hours.

Plating/dilutions

To obtain accurate colony counts of cultures, serial dilution and plating were used. Serial dilutions were 10-fold, and used 0.85% (w/v) sterile saline (NaCl). Two methods of plating were used: [1] spread-plates, where 50 or 100 µl dilutions were spread onto LB agar with an L-shaped spreader; [2] 20 µl droplets were used and were allowed to dry before incubation (Miles et al. 1938). In both cases plates were incubated at 37°C overnight.

Optical density readings

Optical density (OD) readings were taken in two ways: using a Nanodrop spectrophotometer (Thermo Scientific, USA), or using a plate reader (SpectraMax 190, Molecular Devices UK Ltd., Wokingham, UK). In both cases OD_{600} of o/n cultures was measured. Nanodrop readings used 1.5 µl culture, the plate reader used 100 µl.

Long term strain storage

Fresh o/n cultures were added to 80% glycerol in a 1:4 ratio and immediately frozen in dry ice. Stock cultures were stored long term at -80°C. Cultures were revived by plating from glycerol stocks onto LB agar (with appropriate antibiotics if required) and incubating o/n.

PCR

Two PCR enzymes were used in this thesis: Taq (Qiagen Ltd, Manchester, UK) and Easy-A cloning enzyme (Agilent Technologies, Wokingham, UK). Here standard PCR conditions are given for both enzymes. Specific annealing temperatures for each primer set are given in the appropriate methods sections. Easy-A standard reagent concentrations: dNTPs 10mM, Taq 2.5 units and primers 0.2 pmol μl^{-1}. Cycling conditions: initial denaturation, 95°C, 2 minutes, followed by 30 cycles of denaturation (95°C, 40 seconds), annealing (30 seconds) and extension (72°C, 60 seconds) and a final extension step of 72°C for 7 minutes. Taq standard reagent concentrations: dNTPs 0.4mM, Taq 0.625 units and primers 0.2 µM. Cycling conditions: initial denaturation, 95°C, 3 minutes, followed by 30 cycles of denaturation (95°C, 30 seconds), annealing (30 seconds) and extension (72°C, 60 seconds) and a final extension, step of 72°C for 3 minutes.

Colony PCR

PCR reactions were prepared as described above except that cells from a single o/n colony were used as the DNA template. A sterile toothpick was used to touch the edge of a single colony from a fresh o/n plate. The cells were transferred to the reaction tube and the stick rapidly swirled for a few seconds to release the

cells. The cycling conditions were as described above, except for the addition of a long initial denaturation step of 95°C for 5 minutes (Woodman 2008).

2.2.2 Production of mutants

The pCT plasmid was isolated from *E. coli* DH10B using a High-speed Midi Kit (Qiagen). The extraction was run on an agarose gel to check the purity, as described by Gonzalez et al. (1982). Electrocompetant DH10B and MG1655 cells were prepared according to the Bio-Rad protocol (Bio-Rad Laboratories, UK): 300ml LB broth was inoculated 1:100 with an 8-hour starter culture of the appropriate *E. coli*, and was then incubated at 37°C, and shaken at 220 rpm, overnight. This culture was then chilled on ice for 20 minutes before harvesting cells by centrifugation at 4000 x *g* for 15 minutes at 4°C. The supernatant was poured off and the pellet resuspended in 300 ml ice-cold 10% (w/v) glycerol, and re-centrifuged at 4000 x *g* for 15 minutes at 4°C. This process was repeated twice more, resuspending first in 150 ml and then in 12 ml ice-cold 10% glycerol. After the final centrifugation step, cells were resuspended into 2 ml ice-cold 10% glycerol, and stored in 200 µl aliquots at -80°C. Electrocompetant cells were transformed with the helper pLGBE plasmid (Bloor & Cranenburgh 2006) and extracted pCT (DH10B only) using the MicroPulser™ electroporation apparatus (Bio-Rad, Hemel Hempstead, UK). Here electrocompetant cells were thawed on ice and 40 µl cell suspension was mixed with 1 µl DNA (approximately 100 ng μl^{-1}) in a chilled microcentrifuge tube and chilled on ice for 1 minute. The DNA/cell mixture was added to an ice-cold 0.1 cm cuvette and electroporated at approximately 1.8 kV for 5 ms. Immediately afterwards 1 ml of SOC media (2% Bacto-tryptone, 0.5% Bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgSO_4$ and 20 mM glucose) was added and cells were transferred to universal tube and incubated at 37°C for 1 hour, before plating on onto LB Agar containing ampicillin and incubating overnight.

Mutants were produced using the Xercise protocol, as described by Bloor and Cranenburgh (2006). This involved PCR amplification of dif-CAT-dif motif from plasmid pTOPO-DifCAT using primers with 5' homology to the deletion target

site (Table 2.2) and Easy-A cloning enzyme (Agilent Technologies, Wokingham, UK) (annealing temperature 65°C). PCR products were then purified (QIAquick gel extraction kit, Qiagen). Host cells (DH10B or MG1655) carrying pLGBE and induced with 1% arabinose were electrotransformed with the purified PCR products and selected with chloramphenicol. The PCR product carrying a chloramphenicol resistance gene homologously recombines at the target deletion site, displacing the unwanted region (see Figure 2.1). Transformant colonies were then transferred to non-selective LB broth culture, and replica plated onto tetracycline, chloramphenicol and LB plates. The helper pLGBE plasmid facilitates site-specific recombination at *dif* sites, excising the chloramphenicol resistance gene. Transformants with successful deletions and subsequent recombination are chlorampenicol sensitive; and the loss of the helper pLGBE plasmid restores tetracycline susceptibility. A second Δ*lac* mutant that retained the CAT (chloramphenicol acetyltransferase) resistance gene was produced by excluding the *dif* sites from the PCR construct.

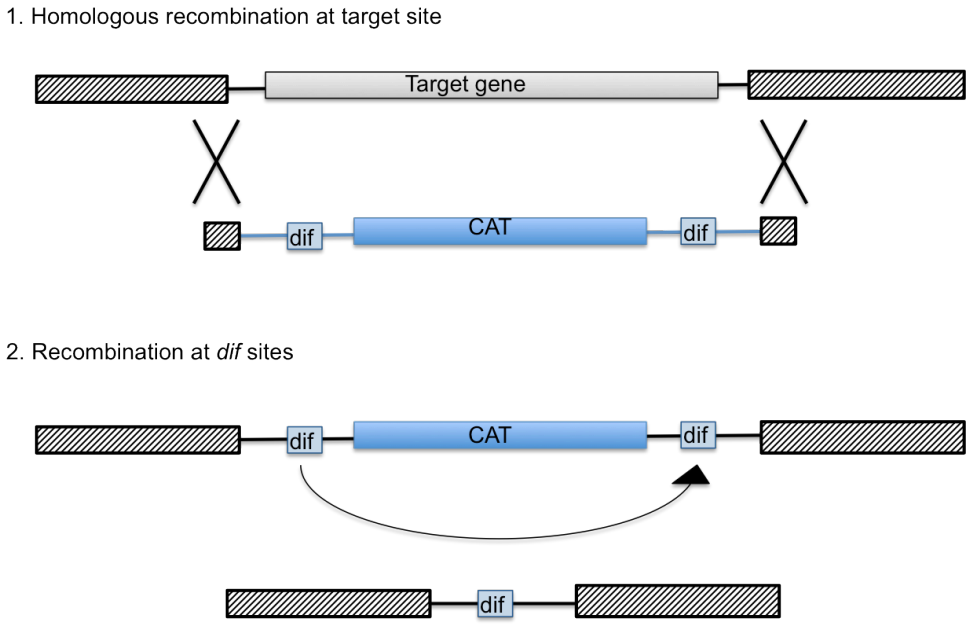

Figure 2.1: Xercise gene deletion method. The dif-CAT-dif motif of the pTOPO-DifCAT plasmid is amplified using primers with homology to the target deletion site. This construct then homologously recombines at the target site to excise the desired gene (Step 1). The dif sites then recombine to excise the selective chloramphenicol resistance marker (Step 2).

Diagnostic PCR was used to assess the insertion and subsequent deletion of the dif-CAT-dif fragment into pCT and MG1655. Diagnostic PCR was conducted using Taq (Qiagen) with an annealing temperature of 55°C. Deletion primers for MG1655 Δlac +/- CAT were taken from Datsenko & Wanner (2000). All other primers were designed using Primer3 (Rozen & Skaletsky 2000) and Serial Cloner 2.5 (Perez 2013) (Table 2.2).

Electrocompetant MG1655, MG1655 *Δlac* and MG1655 *Δlac* CAT$^+$ cells were prepared according to the Bio-Rad protocol, and transformed with both pCT and pCT*ΔtrbA* plasmids using MicroPulser electroporation apparatus (Bio-Rad) as described above.

Name	**Sequence (5'-3')**	**Size (nt)**
F trbA del	TTCTGCATCAACGGTATCAACAAGCACCGTTTCAGT TATTTC**AGTGTGCTGGAATTCGCCCT**	62
R trbA del	ATTGTTCGCATTAATTCCACTCAGCCTCATCCCGAA ATTTATCTATTTAT**CTGCAGAATTCGCCCTTCCT**	70
F trbA diag	CGGCATCCAGGCAGGCATCA	20
R trbA diag	TTCAGCCCTGCCCGGTCATT	20
F lac del	ATTGCGGCCTATATGGATGTTGGAACCGTAAGAGA A**AGTGTGCTGGAATTCGCCCT**	56
R lac del	GCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTA G**CTGCAGAATTCGCCCTTCCT**	56
F Lac del CAT$^+$	ATTGCGGCCTATATGGATGTTGGAACCGTAAGAGA A**CAGGCGTAGCACCAGGCGTT**	56
R Lac del CAT$^+$	GCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTA G**GGGAAGCCCTGGGCCAACTT**	56
F Lac diag	ACGGAAAGAGTAACGTTGGGTGC	23
R Lac diag	GCGCCATTACCGAGTCCGGG	20

Table 2.2: *trbA* and *lacZYA* deletion & diagnostic primers. Primers designated 'del' produced the homologous recombination construct. **Bold**: homology to pTOPO-DifCAT plasmid; Standard: homology to target gene. Primers designated 'diag' were used to diagnose successful insertions/deletions.

2.2.3 Confirmation of mutant genotypes

Diagnostic PCR using primers *trbA diag* or *lac diag* (Table 2.2), identified potentially successful mutants by comparing product sizes on an agarose gel (Voytas 2000). To confirm genotypes of mutant strains these PCR products were sequenced. PCR products were cleaned up using Exo1 (Fisher Scientific) and TSAP (Promega UK, Southampton) both at final concentration 0.045 units μl^{-1}. Clean up mixes were incubated at 37°C for 30 minutes, followed by 80°C for 15 minutes. Sequencing reactions were conducted using BigDye® terminator v3.1 cycle sequencing kit (Life Technologies Ltd. Paisley, UK) and both forward and reverse diagnostic primers. Cycling conditions were as follows: 96°C for 1minute, followed by 25 cycles of: 96°C for 10 seconds, 50°C for 5 seconds, 60°C or 4 minutes. Sequencing reactions were cleaned up using ethanol/sodium acetate precipitation. The sequencing run was conducted by AHVLA Scientific (AHVLA, Weybridge, UK). Sequences were aligned using Geneious version 5.2 (Biomatters 2012) and compared to the wild type sequence.

2.2.4 Confirmation of mutant phenotypes

For the *Δlac* strain first screening was done on LB plates containing X-Gal (0.02 mg ml^{-1}) and IPTG (0.1 mM) (Fisher Scientific), to assess for blue/white colour. White strains were confirmed by sequencing. For the CAT+ strain, chloramphenicol was also added to selective plates.

Mating experiments

To confirm the conjugation null phenotype of the pCT*ΔtrbA* plasmid mating experiments were conducted with both wild type and mutant pCT. Mating experiments protocols were adapted from Liebana et al. (2004): OD_{600} of o/n cultures of donor (MG1655+ pCT/pCT*ΔtrbA*) and recipient (MG1655 *Δlac* CAT^+) strains were measured, and crosses set up from independent cultures. Mixtures were made with a 1:5 ratio of donors to recipients and a 1 in 5 dilution into fresh LB broth, with a final volume of 3ml. Mixes were incubated at 37°C, 220rpm for 48 hours. Single strain control cultures were also incubated. Initial mixtures were serially diluted and plated onto LB + cefotaxime or chloramphenicol to

assess initial ratios. After 48 hours, mixes and single strain cultures were plated onto cefotaxime and mixed antibiotic plates to determine final ratios and numbers of transconjugants, as shown in the formula:

$$\textit{Transconjugants per donor} = \frac{\textit{Total transconjugants}}{\textit{Total donors}}$$

Single strain cultures were used to determine the rate of spontaneous mutation to cefotaxime/chloramphenicol resistance. Two replicates of this experiment were run, with n = 3 and n = 6 for each mix. These data were compiled and analysed together.

Mating experiments were also conducted using a rifampicin resistant *Enterobacter* strain, jjbc, as the recipient. Donors and recipients were again mixed in a ratio of 1:5 and diluted 1:5 in LB broth, with a final volume of 3ml. Cultures were incubated either shaken at 220rpm, or unshaken, for 5 hours at 37°C. Dilutions were then plated onto cefotaxime only and cefotaxime + rifampicin plates; these were incubated at 37°C overnight, and colonies counted.

Competition experiments

The fitness cost of the *lacZYA* knockout was assessed by competition experiments between blue and white strains (MG1655 and MG1655Δ*lac*). Competitions were conducted in LB broth and on LB agar plates, both containing X-Gal & IPTG. OD_{600} of o/n cultures of competing strains were measured and cultures diluted to approximately equal cell density with 0.85% saline. For plate competitions, cultures were then mixed 50:50 and diluted with saline to give an approximate cell density of 1 colony forming unit (CFU) per μl (CFU μl^{-1}). To confirm the initial ratio of cells, mixed cultures were plated onto LB agar + X-Gal & IPTG and incubated o/n at 37°C, and blue and white colonies counted. For the competitions, 100 μl mixed cultures were spread onto plates and incubated at 37°C for 72 hours, after which colonies were harvested by adding 5ml 0.85% saline to the plate, loosening colonies with a sterile spreader and homogenizing with a pipette. Harvested colonies were serially diluted and plated onto LB agar + X-Gal & IPTG, with and without ampicillin. Plates were incubated at 37°C

overnight, and colonies counted to determine final ratios, and relative fitness of the blue (MG1655) strain.

Broth competitions were conducted in 1ml LB broth + X-Gal & IPTG in 24 well plates (Costar® Corning®, Sigma-Aldrich), inoculated 1:5 with diluted strains. 20 µl was immediately taken and diluted to obtain the starting ratio. Broth cultures were incubated at 120 rpm, 37°C for 24 hours, and then serially diluted and plated out onto LB agar + X-Gal & IPTG +/- cefotaxime to get final ratios and calculate relative fitness of the blue strain (MG1655). Relative MG1655 fitness was calculated as described by Lenski (1988), and is given by the equation:

$$W_{ij} = \frac{log_2 N_i(1)/ N_i(0)}{log_2 N_j(1)/N_j(0)}$$

Where W_{ij} is the fitness of strain i (MG1655) relative to strain j (MG1655 *Δlac*), and N is the cell density at the start (0) and end (1) of the experiment.

2.3 Results

2.3.1 Sequencing confirms mutant genotypes

Diagnostic PCR using primers *trbA diag* or *lac diag* (Table 2.2), identified mutants with successful insertions and subsequent deletion of the dif-CAT-dif motif (dCd) in the *trbA* and *lacZYA* target sites respectively. Figure 2.2 shows an agarose gel of this diagnostic PCR for the *trbA* site. Lane 5 shows wild type pCT, lanes 3 & 4 show insertion of the dCd motif, and lanes 1 & 2 show deletion of dCd by Xer recombinase following removal of chloramphenicol selection. The predicted fragment sizes are: wild type pCT 1579bp; insertion 1182bp; deletion 269bp. For the *lacZYA* mutant, predicted product sizes are: wild type MG1655 5222bp; insertion 1563bp; deletion 537bp. *lacZYA* CAT$^+$ mutant: wild type MG1655 5222bp; insertion 1323bp. Mutants with appropriate PCR product sizes were sequenced to confirm the deletion. Alignments of sequences from forward and reverse primers and from different mutants were carried out in Geneious version 5.2 (Biomatters 2012). The alignments confirmed that the deletions had occurred in the correct genes (Figure 2.3 shows the alignment for pCT).

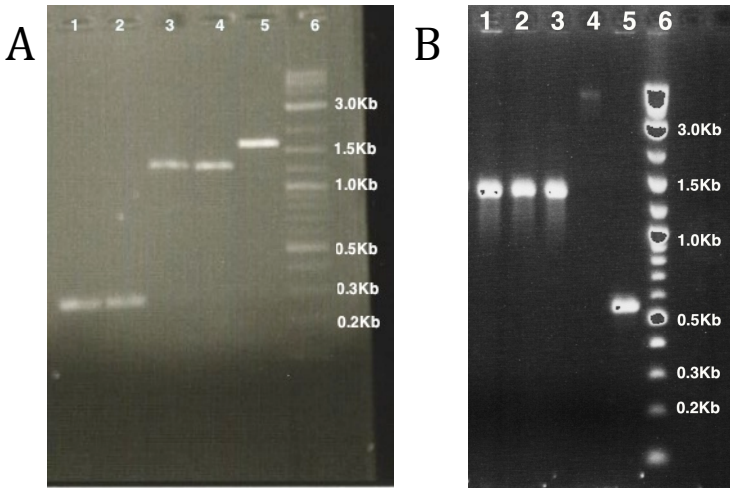

Figure 2.2: Gels showing PCR products of diagnostic PCR for pCT*ΔtrbA* (A) and MG1555 mutants (B). A Lane 5: wild type pCT; lanes 3 & 4: insertion; lanes 1 & 2: deletion of dCd; Lane 6: 2-log ladder (NEB). Predicted fragment sizes: wild type: 1579bp; insertion: 1182bp; deletion: 526bp. B Lanes 1-3: *ΔlacZYA* CAT$^+$ insertion; lane 4: wild type pCT; lane 5: *ΔlacZYA* dCd deletion; Lane 6: 2-log ladder (NEB). Predicted fragment sizes: wild type MG1655 5222bp; *lacZYA* CAT$^+$ insertion 1323bp; *ΔlacZYA* dCd deletion 537bp.

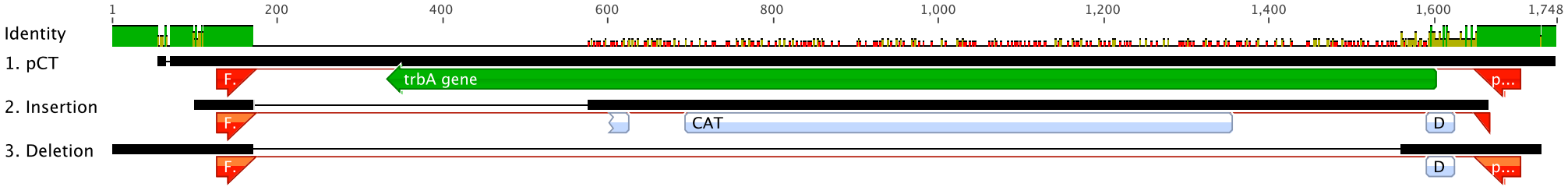

Figure 2.3: Alignment of original pCT plasmid *trbA* region with dif-CAT-dif insertion and dif-recombined deletion sequencing products. The top line shows the % identity between the sequences (n 100% identity). Sequences are annotated with the dif-CAT-dif sites (n) and the homology region primer sites (n). Large gaps (—) were inserted into the insertion and deletion sequences to allow alignment of the primer sites.

2.3.2 ΔtrbA mutants are non-conjugative

Mating experiments with the MG1655 *Δlac* CAT$^+$ recipient demonstrated almost no conjugation in the Δ*trbA* mutant compared to the wild type pCT. Rates of conjugation per donor were significantly lower for the mutant plasmid: 4.43 x 10^{-4} for pCT and 1.10 x 10^{-6} for Δ*trbA* (p = 6.1 x 10^{-8}) (Figure 2.4). In fact, in 6 out of 9 Δ*trbA* crosses no transconjugants were detected at all.

Mating experiment with rifampicin resistant *Enterobacter* as the recipient also showed significantly less conjugation in the Δ*trbA* mutant. The mean number of transconjugants per donor with pCT was 2.13 x 10^{-6}, whereas for the Δ*trbA* mutant it was 3.31 x 10^{-7} (p = 0.047). However, rates of conjugation were lower with the *Enterobacter* recipient, and rates of spontaneous mutation were higher (data not shown), therefore *Enterobacter* was determined to be a less useful recipient than the CAT$^+$ marked MG1655.

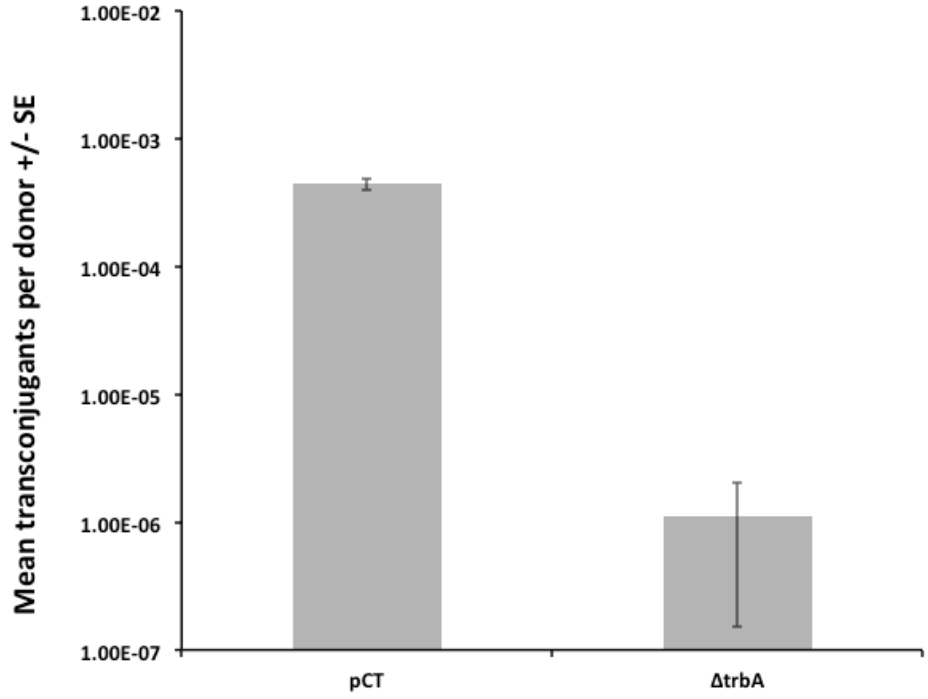

Figure 2.4: Mating experiments with wild type and Δ*trbA* pCT. The mutant plasmid has a significantly lower conjugation rate than the wild type (p= 6.1 x 10^{-8}).

2.3.3 ΔlacZYA mutants do not metabolise X-Gal & IPTG

Screening on X-Gal & IPTG plates showed that *ΔlacZYA* mutants did not break down X-Gal, and colonies remained white. Figure 2.5 shows the relative fitness (*W*) of the blue MG1655 strain in competition with the white *ΔlacZYA* knockout on LB agar plates and in LB broth. In plate competition, mean relative MG1655 fitness is 0.978 (95% confidence limits: 0.96 – 0.99), indicating that the blue strain does very slightly less well than if both strains were of equal fitness, that is $W = 1$ ($p = 2.50 \times 10^{-6}$). Although this difference is significant, it is very small and unlikely to have a huge impact on further competition experiments. In contrast, despite the relative fitness of MG1655 being higher in broth (mean = 1.02, 95% confidence limits: 0.72 – 1.31), this is not a significant difference from equal fitness ($p = 0.47$). Exclusion of the outlier, $W = 1.92$, increases the mean to 1.37 (95% confidence limits: 1.22 – 1.51). Although this difference is still not significant, $p = 0.13$, this may be due to low replication in this test (n = 5 excluding the outlier).

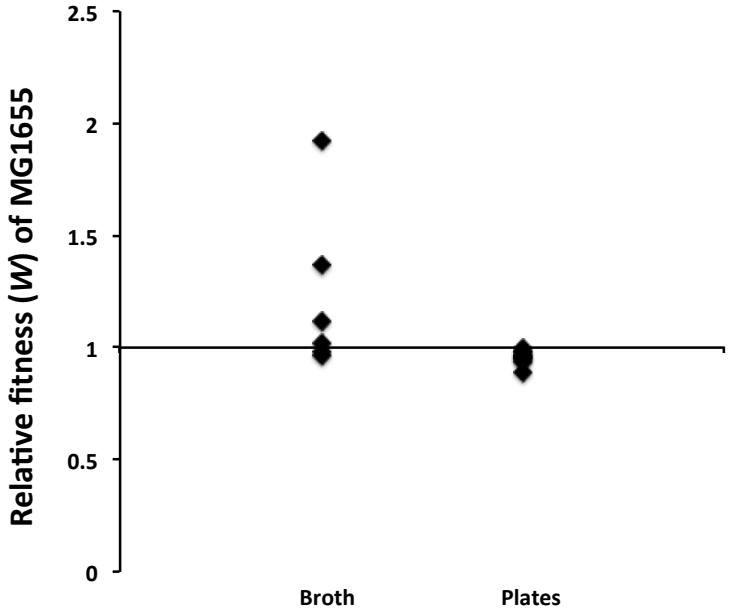

Figure 2.5: Relative fitness of MG1655 in competition with Δ*lacZYA* knockout. In plate competition MG1655 has slightly lower relative fitness (mean = 0.978, 95% confidence limits: 0.96 – 0.99, n = 10) whereas in broth it is slightly higher (mean = 1.02, 95% confidence limits: 0.72 – 1.31, n = 6). Relative MG1655 fitness (*W*) calculated as described by Lenski (1988). *W* >1 indicates higher fitness relative than the competitor

2.4 Discussion

The Xercise gene deletion method (Bloor & Cranenburgh 2006) was successful in producing both the non-conjugative pCTΔ*trbA* mutants (CAT +/-) and the white MG1655 Δ*lacZYA* mutant, as demonstrated by both genotypic and phenotypic assays. Sequencing of the target regions showed correct insertion and subsequent recombination at the *dif* sites for all mutants. Phenotypic assays demonstrated that the deletions had the desired effects. Mating experiments with pCTΔ*trbA* showed significantly lower conjugation rates than wild type pCT with two different recipient strains, *E. coli* MG1655 and *Enterobacter*. *Enterobacter* was determined to be a less useful recipient as rates of conjugation where considerably lower, even with the wild type plasmid, therefore MG1655 pairs are used in all subsequent experiments.

The MG1655 Δ*lac* mutant does not break down X-Gal and colonies remain white, whereas standard MG1655 colonies are blue. Fitness assays suggest that although there are differences between the two strains on plates and in broth, these differences are small and unlikely to affect the outcome of further experiments. However, these results were taken into consideration in the analysis of subsequent competition experiments with these strains.

References

Bio-Rad, High efficiency electrotransformation of *E. coli*, Bio-Rad Laboratories UK.

Biomatters, 2012. Geneious version 5.2. Available at: http://www.geneious.com/.

Bloor, A. & Cranenburgh, R., 2006. An efficient method of selectable marker gene excision by Xer recombination for gene replacement in bacterial chromosomes. *Applied and Environmental Microbiology*, 72(4), pp.2520–2525.

Cottell, J.L. et al., 2011a. Complete sequence and molecular epidemiology of IncK epidemic plasmid encoding blaCTX-M-14. *Emerging infectious diseases*, 17(4), pp.645–652.

Cottell, J.L., 2011b. Investigation of factors influencing the successful persistence and dissemination of a globally distributed antibiotic resistance plasmid. PhD Thesis, University of Birmingham, UK.

Datsenko, K. & Wanner, B., 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences*, 97(12), pp.6640–6645.

González, J.M., Brown, B.J. & Carlton, B.C., 1982. Transfer of *Bacillus thuringiensis* plasmids coding for delta-endotoxin among strains of *B. thuringiensis* and *B. cereus*. *Proceedings of the National Academy of Sciences of the United States of America*, 79(22), pp.6951–6955.

Komano, T. et al., 2000. The transfer region of IncI1 plasmid R64: similarities between R64 tra and *Legionella* icm/dot genes. *Molecular Microbiology*, 35(6), pp.1348–1359.

Lenski, R., 1988. Experimental studies of pleiotropy and epistasis in *Escherichia coli*. I. Variation in competitive fitness among mutants resistant to virus T4. *Evolution*, 42(3), pp.425–432.

Liebana, E. et al., 2004. Characterization of β-lactamases responsible for resistance to extended-spectrum cephalosporins in *Escherichia coli* and *Salmonella enterica* strains from food-producing animals in the United Kingdom. *Microbial Drug Resistance*, 10(1), pp.1–9.

Liebana, E. et al., 2006. Longitudinal farm study of extended-spectrum β -lactamase-mediated resistance. *Journal of clinical microbiology*, 44(5), pp.1630–1634.

Miles, A.A., Misra, S.S. & Irwin, J.O., 1938. The estimation of the bactericidal power of the blood. *Journal of Hygiene*, 38(06), pp.732–749.

Perez, F., 2013. SerialCloner version 2.5. Available at: http://serialbasics.free.fr/Serial_Cloner.html.

Praszkier et al., 1991. Comparative analysis of the replication regions of IncB, InK and IncZ plasmids. *Journal of Bacteriology*, 173(7), pp. 2393-2397.

Rozen, S. & Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods in molecular biology (Clifton, N.J.)*, 132, pp.365–386.

Teale, C.J. et al. 2005. Extended-spectrum beta-lactamase detected in *E. coli* recovered from calves in Wales. Veterinary records. 156, 186-187.

Voytas, D., 2000. Agarose gel electrophoresis. In F. M. Ausubel, ed. *Current Protocols in Molecular Biology*. John Wiley and Sons, pp. 2.5A.1–2.5A.9.

Woodman, M.E., 2008. Direct PCR of intact bacteria (colony PCR). In *Current Protocols in Microbiology*. John Wiley & Sons, Inc., pp. 9:3D:A.3D.1–A.3D.6.

Chapter 3: Cooperative β-lactam resistance

3.1 Introduction

3.1.1 Antibiotic resistance as a cooperative trait

A cooperative trait is a behaviour by one individual that can benefit another (West et al. 2006). In bacteria, cooperative traits often come in the form of 'public goods' released into the environment and available to all. Many virulence factors, including siderophore production and biofilm formation are cooperative traits. Antibiotic resistance conferred by the enzymatic breakdown of drugs is potentially a cooperative trait as it can detoxify the environment for all cells, and the production of β-lactamases, which cleave and deactivate penicillins, is often cited as a social trait in bacteria (West et al. 2007; West et al. 2006; Brown et al. 2009). The phenomenon of protective clearance of antibiotics by resistant cells is commonly seen by microbiologists in the presence of 'satellite' colonies on transformation plates (Figure 3.1). These non-resistant colonies are able to grow on ampicillin plates where resistant colonies are already established. It seems likely that resistant transformants clear the antibiotic from their immediate vicinity, creating an antibiotic-free space where susceptible 'satellite' colonies can then grow. Previous studies have demonstrated the survival of antibiotic sensitive *Escherichia coli* and *Salmonella sp.* in the presence of resistant strains at high concentrations of antibiotic *in vitro* (Dugatkin et al. 2005; Perlin et al. 2009; Clark et al. 2009) Cross-species protection of susceptible bacteria by β-lactamase producers has also been seen *in vivo* (Brook et al. 1983; Tacking 1954; Hackman & Wilkins 1975).

Clinical studies have suggested that protective clearance is mediated by the release of β-lactamase enzymes into the environment by producing cells (Brook 2004), and packaging of β-lactamases into extracellular vesicles has been demonstrated in *Pseudomonas aeruginosa* (Ciofu et al. 2000). Secretion of enzymes could increase the area of antibiotic clearance, benefitting all cells in

the area, as seen with the secretion of siderophores. Dugatkin et al. (2005) use two engineered plasmids, drawing a distinction between 'self-limited' and 'shared' resistance, based on the localisation of the enzyme to the inner periplasmic or outer membranes respectively. Although secretion of enzymes, either alone or inside vesicles, may facilitate detoxification of a larger area via diffusion, the distinction between 'shared' and 'self-limited' resistance is questionable; is secretion of β-lactamases necessary for detoxification of the medium? The fact that both strains facilitated growth of *Salmonella* suggests that it is not (Perlin et al., 2009). This chapter uses the naturally occurring extended spectrum β-lactamase (ESBL) plasmid, pCT, to investigate the conditions where antibiotic resistance is cooperative.

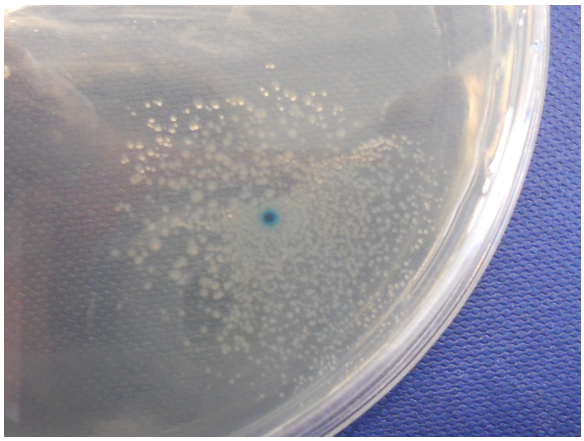

Figure 3.1: Satellite colonies around a successful pUC19 transformant on ampicillin agar. Transformation of *E. coli* DH10B with pUC19 confers resistance to ampicillin and also restores the lac operon, resulting in a blue colony on agar containing X-Gal & IPTG. White susceptible colonies can be seen growing around the resistant transformant, these appeared after >24 hours incubation, whereas successful transformants appeared after ~16 hours.

3.1.2 Persistence and antibiotic tolerance

This chapter initially set out to examine the cooperation in plasmid mediated antibiotic resistance using a naturally occurring resistance plasmid, in order to explore the possibility of 'cheat' plasmids occurring in wild bacterial populations (Chapter 1, Section 1.3.2). However, the results of the initial competition experiments showed no cooperative resistance, which was directly at odds with the results of previous work (Dugatkin et al. 2005; Perlin et al. 2009; Yurtsev et al. 2013), my own bioassay data, and the observation of satellite colonies (Figure 3.1). I therefore turned to the investigation of persisters as an alternative explanation for these results.

Persister cells were identified early in the clinical life of penicillin (Bigger 1944), but a recent resurgence in interest has been fuelled by a wider appreciation of their clinical importance, especially in the light of the current antibiotic resistance crisis (Lewis 2007). Persister cells are natural variants present at low frequency in the bacterial population, which can survive high antibiotic concentrations by dormancy (Lewis 2010). The phenotypic switch between persistence and active growth appears to occur at random, although it is affected by growth phase (Balaban 2004). The presence of persisters in biofilms is thought to contribute to increased antimicrobial tolerance and the maintenance of chronic infections (Brooun et al. 2000; Lewis 2001; Harrison et al. 2005). The second part of this chapter examines the exploitation of resistant β-lactamase producers by persister 'cheats'. The clinical importance of persisters makes their potential evolvement in social cheating all the more interesting.

3.1.3 Aims

The primary aim of this chapter was to examine the cooperative β-lactam resistance using a naturally occurring resistance plasmid, pCT, (Cottell et al. 2011) and to investigate both the environmental and population conditions under which cooperative resistance occurs. To this end, competition experiments were conducted between the pCT-carrying strain and an otherwise isogenic plasmid-free *E. coli* under a variety of conditions. The results of these initial competition experiments showed no cooperative resistance, in contrast to previous work and personal observations using a bioassay. In the second part of this chapter the reason for this discrepancy is explored by testing the hypothesis that persister cells facilitate social cheating. This hypothesis is tested by further competition experiments under a variety of conditions, followed by end-point assays to confirm potential persisters.

3.2 Methods

3.2.1 Strains & culture techniques

The resistant strain is *E. coli* K-12 MG1655 + pCT. The plasmid, pCT, carries the ESBL resistance gene CTX-M-14 (Cottell et al. 2011). The susceptible strain is MG1655 *ΔlacZYA*. The resistant strain is blue in the presence of X-Gal & IPTG whereas the susceptible strain is white. The fitness cost of the *lacZYA* knockout was assessed with competition on plates and a negligible fitness difference was found (Chapter 2, Section 2.3.3). Generic culture techniques and antibiotic specifications can also be seen in Chapter 2, Section 2.2.1. The bioassay was also conducted with *E. coli* DH10B + pCT as the resistant strain, and DH10B as the susceptible strain.

3.2.2 Bioassay

A bioassay was used to assess clearance of antibiotic from solid media. A single resistant colony was streaked onto LB agar containing ampicillin and incubated at 37°C for 48 hours (Figure 3.2 B, blue line). Single colonies of susceptible *E. coli* were then streaked from the central resistant colony to the edge of the plates and onto plates with no resistant colony (Figure 3.2, grey lines). Plates were incubated at 37 °C overnight and growth measured from the central colony outwards. In a further assay the central resistant colony was killed using a chloroform soaked filter disk, applied to the central colony for 10 minutes and then removed and allowed to dry before addition of the susceptible strain.

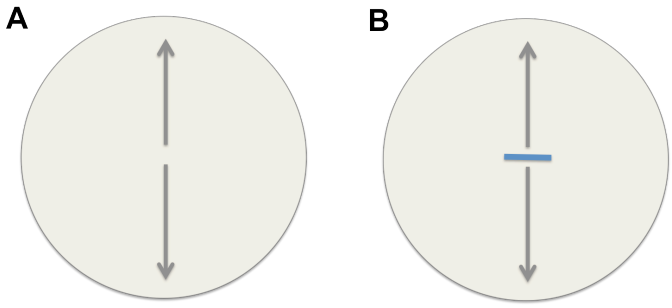

Figure 3.2: Antibiotic clearance bioassay. Ampicillin plates were incubated for 48 hours with (B) or without (A) resistant colonies (blue), before the addition of susceptible colonies (grey). Distance of susceptible growth from central colony was measured after overnight growth.

3.2.3 Minimum inhibitory concentration (MIC) assays

MIC assays were used to determine doses of antibiotics for use in competition experiments, and subsequently to assess changes in antibiotic tolerance in persisters (Andrews 2001). Assays were conducted in 96 well plates (Costar® Corning®, Sigma-Aldrich), excluding the outermost wells to eliminate edge effects cause by evaporation loss. 200 µl LB broth with appropriate antibiotic dose (Table 3.1) was inoculated with 10 µl overnight culture of resistant or susceptible cells or *E. coli* DH10B as a susceptible control. Plates were incubated at 37°C, shaken at 220 rpm, overnight, and the OD_{630} measured with a plate reader (SpectraMax 190). Figure 3.3 shows MIC data for cefotaxime.

Antibiotic	**Dose (µg ml^{-1})**							
Ampicillin	100	20	4	0.8	0.16	0.032	0.0064	0
Cefotaxime	8	1.6	0.32	0.064	0.0128	0.00256	0.00512	0
Nalidixic acid	40	15	10	5	2.5	1	0	-

Table 3.1: Antibiotic doses for MIC assays

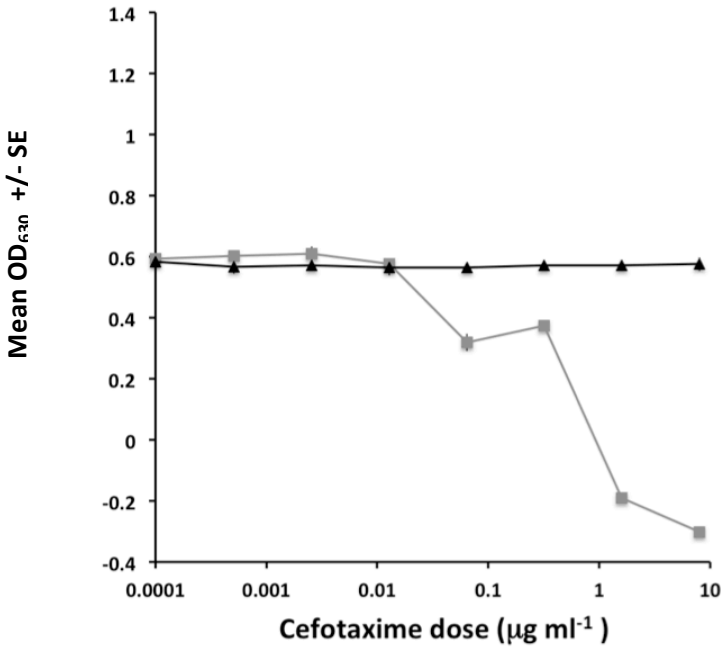

Figure 3.3 Cefotaxime MIC assay. Optical density (OD_{630}) of resistant (MG1655 + pCT, ▲) and susceptible (MG1655, ■) cultures (n = 2) over a range of cefotaxime doses.

3.2.4 Competition experiments

Competitions were conducted on LB agar plates containing X-Gal (0.02 mg ml^{-1}) and IPTG (0.1 mM). OD_{600} of o/n cultures of competing strains (susceptible and resistant) were measured and cultures diluted to approximately equal cell density with 0.85% (w/v) saline. Cultures were then mixed 50:50 and diluted with saline to give an approximate cell density of 1 colony forming unit (CFU) per μl (CFU μl^{-1}). To confirm the initial ratio of cells, mixed cultures were plated onto LB agar + X-Gal & IPTG and incubated o/n at 37°C, and blue and white colonies counted.

For the competitions, plates containing low, intermediate and high doses (Table 3.2) were selected based on the MIC assay (Section 3.2.3, above). 100 μl mixture or susceptible-only control was spread onto plates and incubated at 37°C for 72 hours. Colonies were then harvested by adding 5 ml saline to the plate, loosening colonies with a sterile spreader and homogenizing with a pipette. Harvested colonies were serially diluted and plated onto LB agar + X-Gal & IPTG, with and without ampicillin. Plates were incubated at 37°C overnight, and colonies counted to determine final ratios and relative fitness of the susceptible (MG1655 Δ*lacZYA*) strain (see Section 3.2.7 below).

Four variations of this competition were conducted, altering (i) the initial proportion of resistants (0.1, 0.5, 0.9); (ii) the density of colonies per plate (100, 1000, 10000 CFU); (iii) the nutrient availability (M9 minimal salts 2% agar, Sigma Aldrich); (iv) the antibiotic (cefotaxime, Table 3.2).

Antibiotic	Control	Low	Intermediate	High
Ampicillin	0	4	32	100
Cefotaxime	0	0.064	0.32	8

Table 3.2: Antibiotic doses (μg ml^{-1}) for competition experiments.

3.2.5 Persister competition experiments

In order to increase persister numbers to detectable levels, a large volume of susceptible culture was used to inoculate competition plates. An 8-hour starter culture in 5 ml LB broth was used to initiate 300 ml overnight culture by 1:1000 dilution. The overnight culture was centrifuged at 4000x *g* for 10 minutes, resuspended in 6 ml 0.85% saline and diluted 1:10 (no dilution was used in the varying dose experiment). Mixtures were made 50:50 with an overnight culture of resistant cells. Initial counts and competition plates were set up as described above, using both mixtures and susceptible only cultures (Section 3.2.4). After 24 hours, competition plates were moved to room temperature incubation (approximately 21°C) for up to 3 weeks. Blue and white colonies were counted after 24 hours, and then periodically. After 3 weeks incubation plates were photographed and potential persister colonies screened. Variations were conducted altering (i) the initial number of resistant colonies per plate (2, 10, 100 CFU); and (ii) the ampicillin dose (0, 10, 32, 100 μg ml^{-1}).

3.2.6 Persister confirmation

Ampicillin sensitivity

A subset of potential persister colonies was picked onto LB agar with and without ampicillin, to check for resistance. Persisters were also screened for increased tolerance to ampicillin with an MIC assay (see Section 3.2.3). Persister colonies, plus resistant and susceptible controls, were resuspended in 500 μl saline, and 10 μl of this used as the inoculant.

A growth curve was conducted at a sub-MIC dose of ampicillin, 2 μg ml^{-1}, to assess any change in tolerance. Again persister colonies, plus resistant and susceptible controls, were resuspended in 500 μl saline, and 10 μl of this used to inoculate 200 μl ampicillin broth in a 96 well plate. Plates were incubated in the plate reader at 37°C, shaken, for 10 hours and OD_{595} measured every 15 minutes.

Presence of pCT

To elimate transconjugation as the cause of susceptible survival, a number of persister colonies were PCR screened for the presence of pCT. The pCT specific trbA diag primers (Chapter 2, Table 2.2) were used. Colony PCR protocols are described in Chapter 2, Section 2.2.1.

Antibiotic effect

Low-level ampicillin resistance selected for by continued antibiotic exposure, rather than persistence, could explain the susceptible cells apparent ampicillin tolerance. In order to eliminate this explanation, persister cells were selected by plating susceptible cells onto nalidixic acid. Susceptible cultures were prepared as previously (Section 3.2.5) and then spread onto 15 μg ml^{-1} nalidixic acid plates and incubated at 37°C o/n, followed by room temperature for 6 days. The susceptible strain (MG1655 *ΔlacZYA*) is not resistant to nalidixic acid, and therefore should not grow. Any colonies that emerged on the nalidixic acid plates were taken to be either spontaneous resistance mutants or persisters. Spontaneous resistant mutants appeared after overnight incubation, and were confirmed by re-plating. Colonies that emerged after this (from 24 hours) were taken to be persisters, which had survived the antibiotic dose by dormancy. These persister colonies were screened for any change in ampicillin or nalidixic acid tolerance by MIC assay (as described in Section 3.2.3, for doses see Table 3.1).

3.2.7 Data analysis

Relative susceptible fitness in competition experiments was calculated as described by Lenski (1988; also see Dahlberg & Chao 2003), and is given by the equation:

$$W_{ij} = \frac{log_2 N_i(1)/ N_i(0)}{log_2 N_j(1)/N_j(0)}$$

Where W_{ij} is the fitness of strain *i* (susceptible) relative to strain *j* (resistant) and *N* is the cell density at the start (0) and end (1) of the experiment. Initial cell

density (0) was derived from the mean susceptible and resistant counts (CFU ml^{-1}) in the 0 μg ml^{-1} dose for each experiment. It is not possible to calculate relative fitness when there is no detectable growth; therefore all final cell counts were transformed by addition of the minimal detectable value, calculated by the equation:

$$Minimal\ detectable\ value = \frac{Minimum\ possible\ CFU\ count}{Maximum\ possible\ CFU\ count}$$

Relative susceptible fitness across treatments was compared using ANOVA, conducted in *R* (R Core Team 2013). Persister data was analysed using ANOVA and linear regression as appropriate.

3.3 Results

3.3.1 Bioassay demonstrates protective clearance by resistant bacteria

This bioassay provides a clear visualisation of the resistant strain's ability to facilitate growth of susceptible bacteria (Figure 3.4). The sensitive DH10B strain only grows on ampicillin plates in the region around the resistant DH10B+ pCT colony, despite being inoculated to the edges of the plate (bottom row), and does not grow at all in the absence of resistant colonies (top row). Similar results were obtained when the central colonies were killed with chloroform, indicating that susceptible growth is not due to some interaction (such as conjugation) between resistant and susceptible strains (data not shown). The radius of the growth zone can be used as an estimate of the area of antibiotic clearance by the resistant colonies.

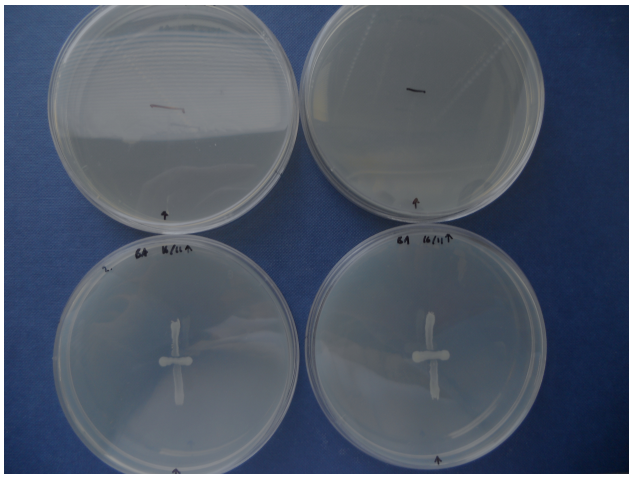

Figure 3.4: Antibiotic clearance bioassay. Susceptible colonies grow on ampicillin plates in the presence of resistant colonies (bottom row), but not alone (top row). Resistant colonies were incubated overnight on 100 μg ml^{-1} ampicillin plates before the addition of susceptible colonies. Susceptible strains can grow in the region around the resistant colony, but do not grow outside the central zone or in the absence of a resistant colony (top row).

3.3.2 Susceptible and resistant cells have equal fitness at sub-inhibitory doses

At sub-inhibitory ampicillin doses (0 & 4 μg ml^{-1}) the relative susceptible fitness is close to 1 across all treatments. It is slightly greater than 1 at 0 μg ml^{-1}, which may indicate higher susceptible fitness in the absence of antibiotic, however the difference is very small. Previous work showed that MG1655 has slightly lower fitness in plate competition than the mutant MG1655 *Δlac* in the absence of antibiotic (mean relative fitness: 0.978, 95% confidence limits: 0.96 – 0.99, see Chapter 2, section 2.3.3). The fact that the susceptible strain has slightly higher relative fitness in these competitions may suggest a fitness cost of plasmid carriage at sub-inhibitory antibiotic doses (Figure 3.5 A).

3.3.3 Relative susceptible fitness declines at inhibitory ampicillin doses

In all competition experiments relative susceptible fitness declines dramatically at inhibitory doses (32 & 100 μg ml^{-1}), with close to zero detectable growth in all cases (Figure 3.5). This result is consistent, regardless of the initial resistant proportion or the density of colonies per plate, indicating that susceptible cells do not benefit from the presence of resistant cells at inhibitory doses of antibiotic, contradicting the bioassay result (Figure 3.4). In the susceptible-only controls there was no growth in the presence of inhibitory ampicillin doses (data not shown).

The initial resistant proportion does not significantly affect relative susceptible fitness ($F_{(2, 55)} = 0.267$, $p = 0.767$), but the dose does have a highly significant effect ($F_{(1, 55)} = 102.64$, $p = 3.49 \times 10^{-14}$, Figure 3.5 A). This is due to the huge change in relative susceptible fitness at inhibitory doses: no susceptible growth was detectable in any replicate at either 32 or 100 μg ml^{-1} ampicillin, and moreover there is no significant difference in relative susceptible fitness between the 0 and 4 μg ml^{-1} doses (means 1.057 and 1.019 respectively, $t_{(28)} = 1.241$, $p = 0.225$).

The density of colonies per plate has no significant effect on relative susceptible fitness ($F_{(2, 54)} = 0.21$, $p = 0.811$, Figure 3.5 B), but this experiment again shows a highly significant effect of antibiotic dose ($F_{(1, 54)} = 62.23$, $p = 1.49 \times 10^{-10}$).

Relative susceptible fitness at inhibitory doses was higher than in the previous experiment and, interestingly, susceptible colonies were detected here. Susceptible colonies were detected in the 100 and 1000 CFU/plate treatments at 32 μg ml^{-1}, and in the 1000 and 10000 CFU/plate treatments at 100 μg ml^{-1}. The higher relative susceptible fitness in the 100 CFU/plate treatment at 32 μg ml^{-1} is likely to be skewed by the low number of replicates (n = 3), two of which had no detectable susceptible growth. This variation is demonstrated by the 95% confidence limits: -0.3663 - 1.6631 (mean: 0.6484). It seems likely that this higher relative fitness is an artefact of low replication, particularly as this treatment is identical to the 0.5 resistant proportion in the previous experiment, where relative susceptible fitness at 32 μg ml^{-1} was -0.8623 with 95% confidence limits of -0.8656 and -0.8591.

More interesting results were obtained from the 1000 CFU/plate treatment at 32 μg ml^{-1} and the 1000 and 10000 CFU/plate treatments at 100 μg ml^{-1}. In the 1000 CFU/plate treatment at 32 μg ml^{-1}, three out of five replicates had detectable susceptible colonies (mean relative susceptible fitness: 0.6409, 95% confidence limits: -0.1212 - 1.4030). At 100 μg ml^{-1} the 1000 CFU/plate treatment just one replicate had detectable susceptible growth (mean relative susceptible fitness: 0.5596, 95% confidence limits: -0.0661 - 1.185), and the 10000 CFU/plate treatment had 2 replicates with detectable susceptible growth (mean relative susceptible fitness: 0.6300, 95% confidence limits: -0.1633 - 1.4232). Although confidence intervals are large for these treatments, the fact that susceptible colonies were detected at all is noteworthy. Although colony density is not a significant predictor of susceptible fitness, these data appear to indicate that higher cell densities may facilitate some susceptible survival at inhibitory antibiotic doses, although these levels are still low. The addition of a selective marker, such as another antibiotic resistance gene, to susceptible cells would improve the detection rate when very low susceptible numbers are present.

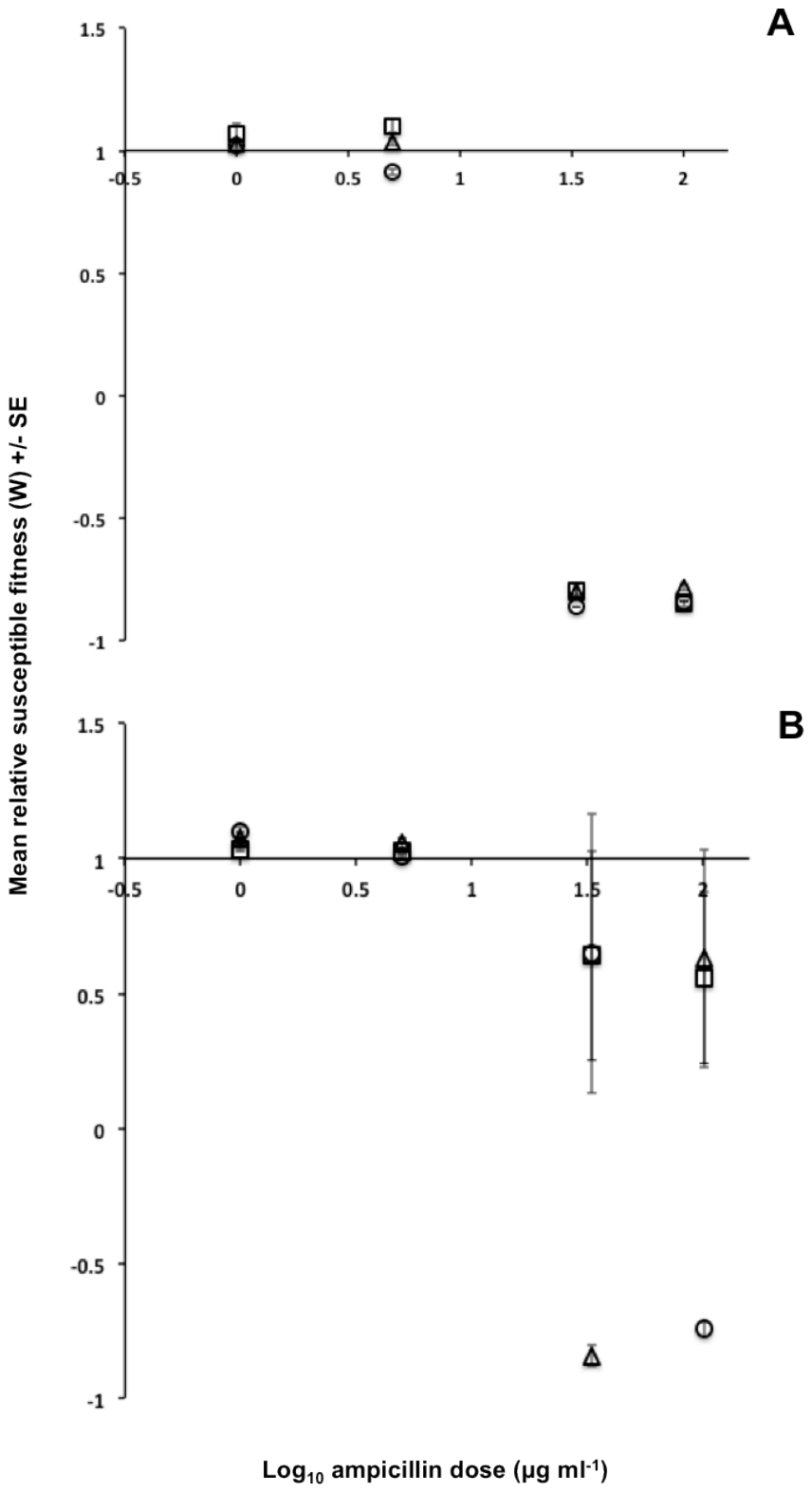

Figure 3.5: Relative fitness of the susceptible strain (MG1655 *ΔlacZYA*) in mixed populations with the resistant strain (MG1655 + pCT) at different antibiotic doses. (**A**) varying initial proportion of resistant bacteria (squares 0.1, circles 0.5, triangles 0.9). (**B**) varying initial cell density (circles 100, squares 1000, triangles 10000 CFU/plate). n = 5 for each data point except (**A**) 0.1 resistant at 32 µg ml^{-1} (n = 4), and (**B**) 100 CFU/plate at 32 µg ml^{-1} (n = 3). Relative susceptible fitness (W) was calculated as described by Lenski (1988), a value >1 indicates higher relative fitness than the competitor, and a value <1 indicate lower relative fitness.

3.3.4 Stress conditions may mitigate plasmid cost

When nutrient availability was limited (M9 minimal media), relative susceptible fitness was reduced at 4 μg ml^{-1} (mean 0.5483, 95% confidence limits -0.0891 - 0.8735) compared to 0 μg ml^{-1} (mean 1.0004, confidence limits 0.9867 - 1.0213, Figure 3.6). This is a curious result, which may suggest that a low nutrient environment reduces antibiotic tolerance, contradicting the generally accepted view that slow growing, nutrient depleted cells are more tolerant to antibiotics (Gilbert et al. 1990; McKenney 1997). It is also possible that the lower relative fitness of the susceptible cells at 4 μg ml^{-1} in M9 minimal media is due to increased sensitivity of susceptibles to ampicillin or some additional benefit of the plasmid in nutrient limited conditions. An MIC assay could be used to assess any change in the antibiotic sensitivity of the susceptible strain in nutrient limited conditions (M9 media). In addition, any possible increased benefit of plasmid carriage in minimal media could be assessed by comparing growth rates of plasmid carrying cells in M9 and LB broths using growth curves at varying doses of ampicillin (as described in Section 3.2.6 above) However, as susceptible growth was observed in just one of five replicates, and the 95% confidence interval is large, further replication is also needed to confirm this result. When cefotaxime replaced ampicillin as the antibiotic, no susceptible growth was detected at all in the sub-inhibitory dose (0.064 μg ml^{-1}). Collectively, the M9 and cefotaxime data indicate that in more stressed conditions the overall cost of the plasmid carriage may be reduced, which would increase the overall benefit associated with resistance even at low doses of antibiotic.

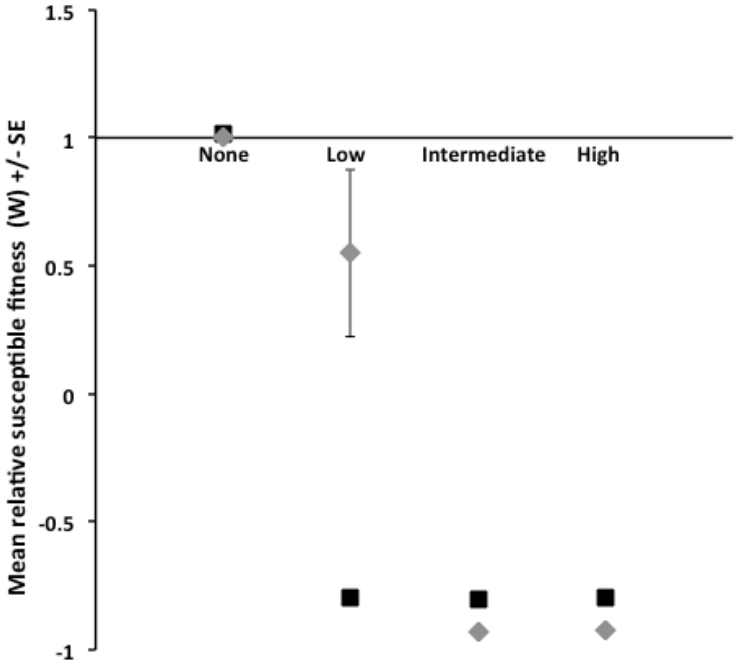

Figure 3.6: Relative fitness of susceptibles in mixed populations in nutrient limited conditions and using cefotaxime. Nutrient limited competitions used M9 media and ampicillin (grey diamonds). Cefotaxime experiments used LB media and cefotaxime instead of ampicillin (black squares). n = 5 for each data point. Relative susceptible fitness (W) was calculated as described by Lenski (1988).

3.3.5 β-lactam resistant bacteria facilitate growth of persisters

The contradiction between the results of the bioassay and the competition experiments can be explained by two factors: the bioassay had higher cell density and it also gave the resistants a head start in clearing the antibiotic. In addition, susceptible colonies were detected at inhibitory ampicillin doses only in the high density treatments. These factors led to the hypothesis that susceptible survival is facilitated by persister cells: natural variants present at low frequency in the susceptible population, which survive high antibiotic concentration by dormancy (Lewis 2010). At high concentrations of susceptibles (3.5×10^8), white susceptible colonies can be observed close to blue resistant colonies even at the highest ampicillin dose, 100 μg ml^{-1} (Figure 3.7). These susceptible colonies were not seen in the absence of resistant bacteria at 100 μg ml^{-1}, but were seen at 32 μg ml^{-1} when double the inoculum volume was used. It

is likely that this susceptible growth is due to degradation of the antibiotic over time (5 days at room temperature), facilitating the growth of persisters. At the higher dose, the antibiotic does not have sufficient time to degrade to a sub-lethal level without the assistance of resistant colonies. This phenomenon was not seen when a lower density of persisters was used (Figure 3.8 B).

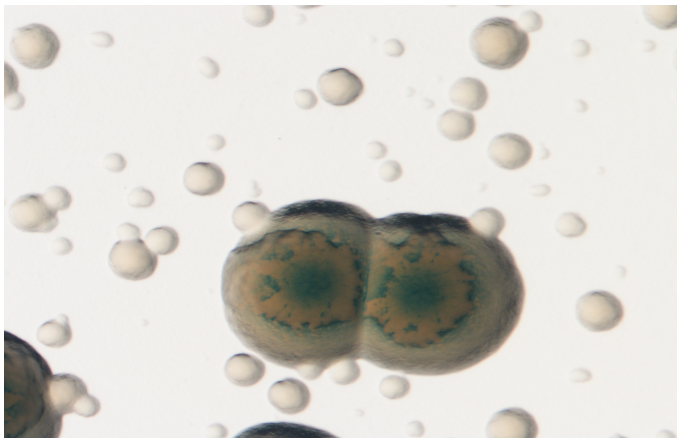

Figure 3.7: Persisters and resistant colonies. x3.0 magnification image showing resistant colonies (blue) and persister colonies (white) on LB agar + 100 μg ml^{-1} ampicillin + X-Gal & IPTG. Resistant colonies appeared after overnight incubation at 37°C. Persister colonies appeared after further 72 hours incubation.

Photographer: Mr K de Souza, Dept. of Earth Sciences, Royal Holloway University of London

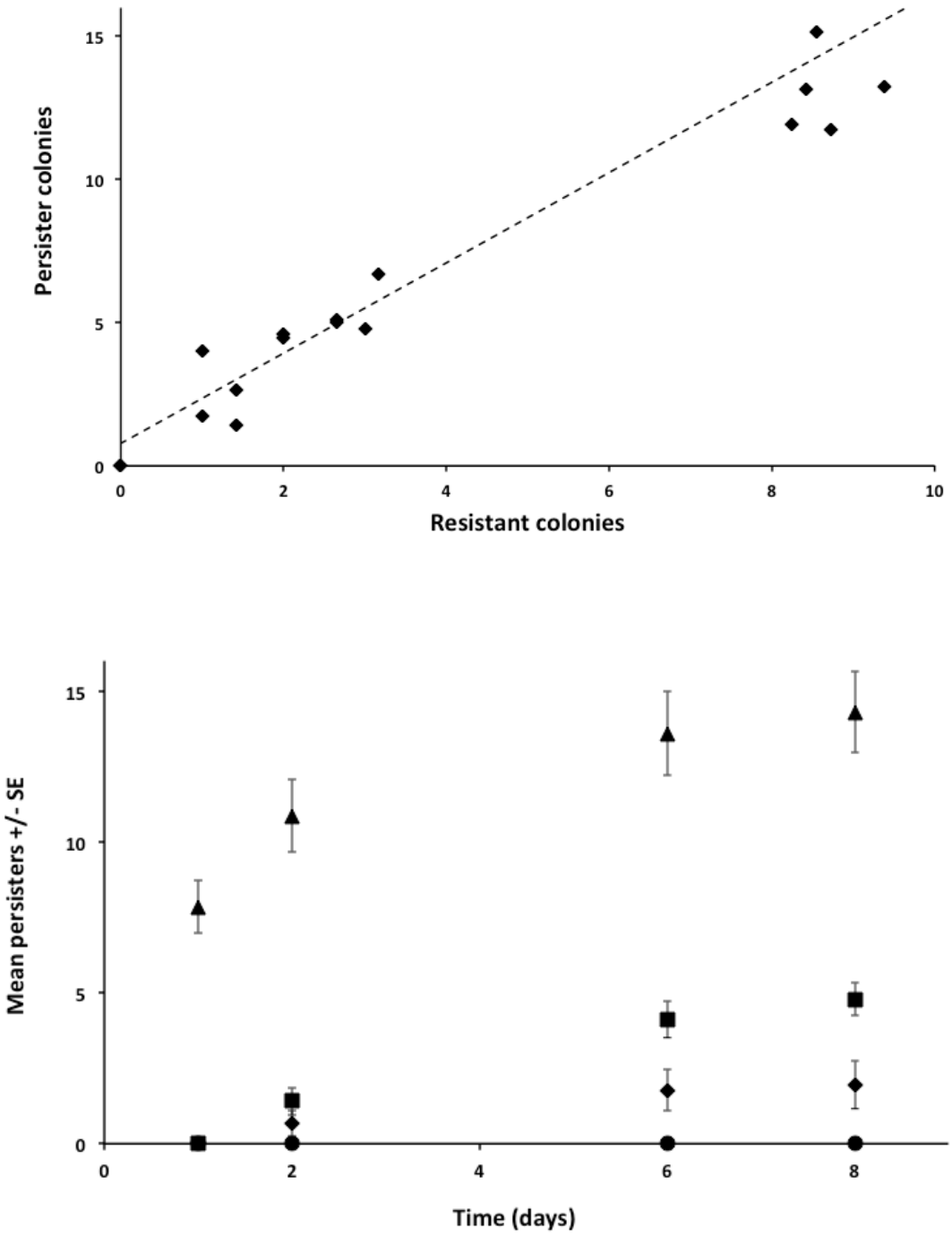

Figure 3.8: The number of persister colonies increases with the number of resistant colonies. (**A**) Numbers of persister and resistant colonies per plate after 8 days incubation (sqrt transformed). Linear regression: Persisters = 0.7801 + resistants x 1.5772. R^2 = 0.8384. $F_{(1,16)}$ = 87.907, p = 6.682 x10^{-8}. (**B**) Mean persister colonies with 0 (●), 2 (◆), 10 (■) or 100 (▲) resistant colonies per plate (n=6, sqrt transformed). There is a highly significant interaction between number of resistant colonies and time, $F_{(1, 92)}$= 17.85, p = 5.62 x 10^{-5}.

The number of these susceptible persister colonies increases with the number of resistant colonies present (Figure 3.8 A; R^2 = 0.8384, $F_{(1,16)}$ = 87.907, p = 6.682

x10^{-8}) and over time, as shown by the highly significant resistant colonies by time interaction term (Figure 3.8 B; $F_{(1, 92)}$= 17.85, p = 5.62 x 10^{-5}). No persisters were found at 100 μg ml^{-1} in the absence of resistant colonies. The number of persisters decreases with increasing antibiotic dose (Figure 3.9; $F_{(1, 50)}$ = 55.62, p = 1.16 x 10^{-9}), and this experiment also shows a significant dose by resistant colony interaction term ($F_{(1, 50)}$ = 6.780, p = 0.0121). These results are consistent with the hypothesis that resistant cells are breaking down the antibiotic in their vicinity, facilitating persister growth: more resistant cells and lower antibiotic concentrations will lead to a more rapid reduction in antibiotic concentration.

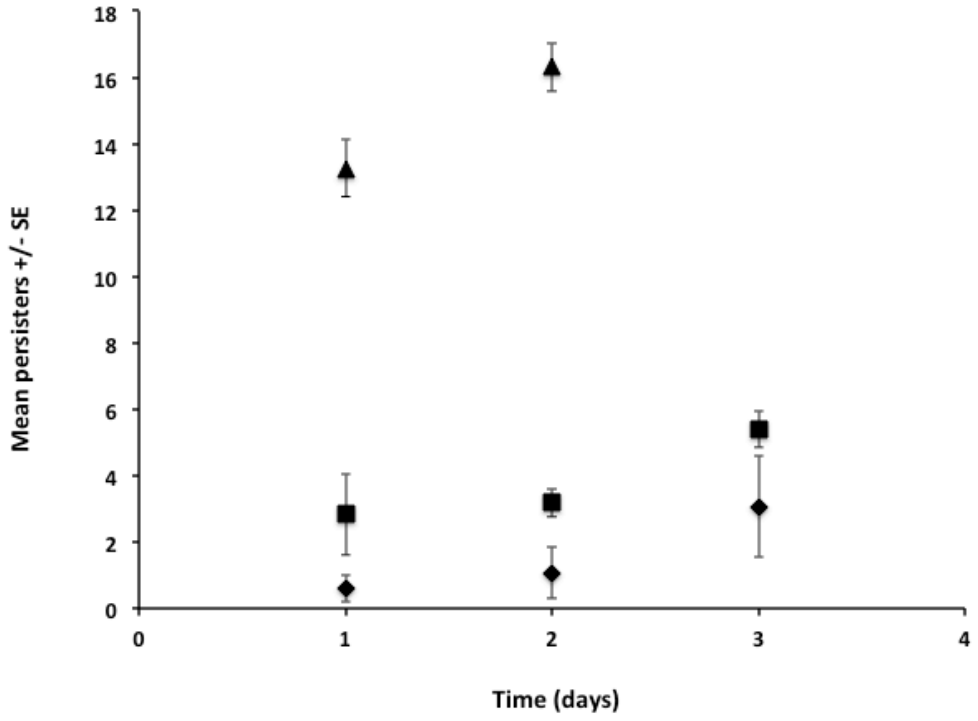

Figure 3.9: The number of persisters decreases with increasing antibiotic dose. Mean persister colonies over time at a variety of ampicillin doses: 100 (◆), 32 (■) and 10 μg ml^{-1} (▲) with ~20 resistant colonies per plate. Data not shown for 10 μg ml^{-1} after 3 days as colonies were too numerous to count (n=7, sqrt transformed). Dose has a highly significant effect on number of persisters $F_{(1, 50)}$ = 55.62, p = 1.16 x 10^{-9}.

3.3.6 Persisters are neither transconjugants nor spontaneous resistance mutants

Persister colonies did not grow when re-plated onto ampicillin agar, indicating that they were not resistant to ampicillin, either by spontaneous mutation or conjugation with resistant cells; nor did PCR using pCT specific primers amplify products (Figure 3.10).

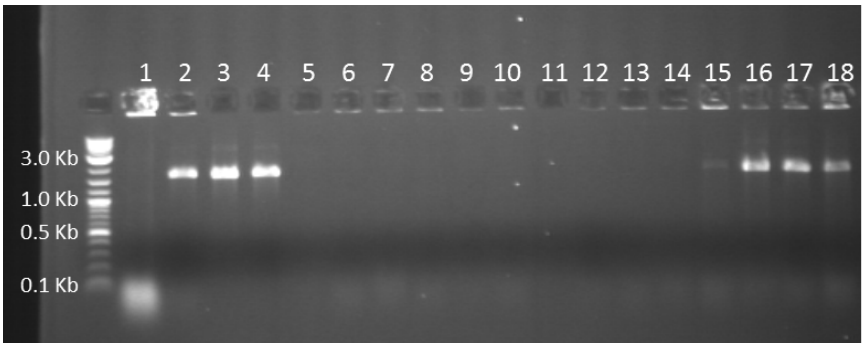

Figure 3.10: Persisters do not amplify pCT specific products. PCR with pCT specific primers detected no pCT in persister colonies. Lane 1: pCT free control (MG1655 *ΔlacZYA*); Lanes 2-4: pCT positive controls (MG1655 + pCT and two pCT plasmid extracts); Lanes 5-14: persister colonies; Lanes 15-18: resistant colonies.

Three further assays were conducted to examine the phenotypes of potential persisters. These were not fully replicated experiments, but provide an indication of persister ampicillin susceptibility in comparison with resistant and susceptible strains. The first assays were two ampicillin MIC tests (Figure 3.11 A & B) for persisters selected on ampicillin and nalidixic acid. The nalidixic acid selected persisters had no prior exposure to ampicillin, in order to rule out any selection for ampicillin tolerance. Persisters showed no difference in ampicillin tolerance compared to susceptibles, whether selected on ampicillin or nalidixic acid. The second assay was a growth curve at a sub-lethal ampicillin dose, 2 μg ml^{-1} (Figure 3.11 C), which indicates that persisters may have slightly increased ampicillin tolerance. Increased tolerance of low doses of antibiotic in cells previously exposed to the same drug has been demonstrated by Adam et al. (2008), who suggest that this is an epigenetic effect. Further data is required to confirm this result, including replication with nalidixic acid selected persisters with no previous exposure to ampicillin.

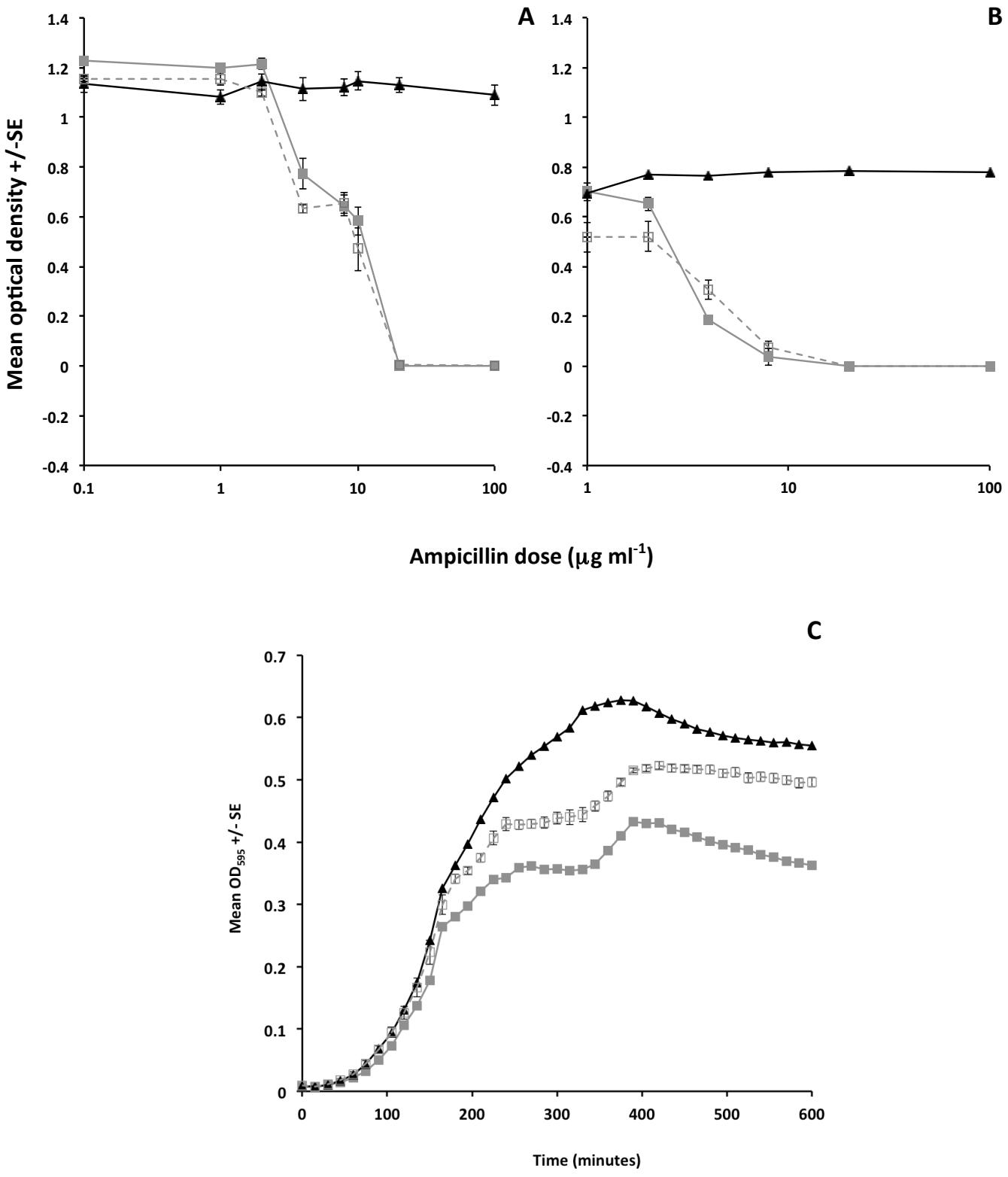

Figure 3.11: **Persister ampicillin susceptibility assays.** (**A & B**) Ampicillin MIC assays for resistant (▲), susceptible (■) and persister cells (□). (**A**) persisters selected on ampicillin (n = 10); (**B**) persisters selected on nalidixic acid (persisters n = 19, susceptibles n = 3, resistant n = 1). (**C**) Growth curves over 10 hours in LB broth with 2 μg ml^{-1} ampicillin (persisters n = 7, resistant and susceptible cells n = 1).

3.4 Discussion

3.4.1 The cost of plasmid carriage

The results of competition experiments in low nutrient conditions and using cefotaxime are curious, as they suggest that stressed conditions reduce the ability of susceptible bacteria to compete with resistant bacteria at sub-MIC doses. This appears to counter the generally accepted view that slow growing, nutrient depleted cells are more tolerant to antibiotics, as many drugs target growing cells, leaving non-growing or very slow growing cells, including persisters, largely unaffected (Hobby et al. 1942; Tuomanen et al. 1986; Gilbert et al. 1990; Cozens et al. 1986). Although it has also been shown that susceptibility can decrease with increased growth rate, this study highlighted the critical effect of environmental conditions on this response (McKenney 1997). Rate of killing has been shown to be proportional to growth rate (Tuomanen et al. 1986), which should give an advantage to susceptible cells in a nutrient depleted environment, as a slow growth rate will reduce killing and allow more cells to survive until the antibiotic is cleared by their resistant neighbours. However, this is not what I have observed. Instead, susceptibles are less competitive in the low nutrient environment at sub-MIC doses. I suggest that the lower growth rate of the susceptibles in nutrient limited conditions may somewhat mitigate the detrimental effects of plasmid carriage on the resistant strain, thus reducing the relative fitness of susceptible bacteria. Dugatkin et al. (2005) and Perlin et al. (2009) demonstrate the cost associated with plasmid carriage and resistance, which can be exploited by susceptible strains, which do not carry either the plasmid or the resistance gene. Similar results were seen using cefotaxime instead of ampicillin. Cefotaxime has been shown to have remarkable potency against Enterobacteriaceae, producing an bactericidal effect at MIC doses (Jones & Thornsberry 1982; Neu 1982). I suggest that a similar mechanism to the nutrient limited model is at work here, meaning that the cost of plasmid carriage is again mitigated by reduced susceptible fitness, even at sub-MIC doses.

3.4.2 Sociality of resistance is dependent on persisters

The resistant strain shows protective clearance of β-lactams, facilitating the growth of susceptible *E. coli*, only under very specific conditions. I posit that these susceptibles are 'persister' cells: natural variants present at low frequency in the susceptible population, which survive high antibiotic concentration by dormancy (Lewis 2010). Very high densities of susceptible bacteria in my experiments increase the numbers of persisters to a detectable level. Once the level of antibiotic is sufficiently reduced by the resistant cells, they begin to appear. This explains both the lag before susceptibles appear, and the increase in numbers of persisters over time, as they switch back to metabolic activity at random (Balaban 2004). Persister cells survive the initial high concentration of antibiotic by remaining dormant, but crucially their ability to grow is a social trait as it is dependent on the frequency of neighbouring β-lactamase producers. Persistence, therefore, facilitates cheating and the exploitation of antibiotic free space provided by the resistant cells.

3.4.3 β-lactams may have variable effects

The observation that there is no social benefit for the susceptible strain except in the presence of persisters contrasts with previous work demonstrating protective clearance of antibiotics by resistant strains (Dugatkin et al. 2005; Perlin et al. 2009; Clark et al. 2009; Yurtsev et al. 2013). One possible explanation for this difference is that different strains and experimental set-ups may produce altered antibiotic tolerance. In this experiment I show that any dose above the MIC is bactericidal to all growing *E. coli.* The bactericidal action of β-lactams is known to be variable and dependent on the specific drug, the strain, growth rate and stage, and the nutrient availability (Hobby et al. 1942) (Tuomanen et al. 1986) (Rolinson et al. 1977; Cozens et al. 1986). It is therefore possible that the experimental design of the earlier papers facilitated susceptible tolerance of doses above MIC, either through strain selection or some element of the growth conditions, producing a bacteriostatic response to ampicillin rather than a bactericidal effect. This tolerance effect has been seen for mercury

resistance, where a threshold exists above which susceptible cells are inhibited but not killed (Ellis et al. 2007).

3.4.4 Pre-clearance of antibiotic required for susceptible growth

Perhaps a more likely explanation is that the part-batch, part-chemostat conditions used in these previous studies facilitate more rapid and generalised clearance of antibiotic than the strict batch culture method I have employed. The model used by Dugatkin et al. (2005) and Perlin et al. (2009) allows build up of β-lactamases in one half of the chamber, and periodically replenishes the antibiotic in the other half. This build up of enzyme facilitates rapid breakdown of the antibiotic, similar to the earlier chemostat model described by Levin (1988). Here coexistence occurs only when susceptibles are added to an already established resistant culture (carrying the degradative chloramphenicol acetyl transferase gene). In a subsequent study, Dugatkin and colleagues recognise that enzyme accumulation may be a contributory factor, and therefore exchange between the two sides of the chamber and the frequency of antibiotic addition are increased (Clark et al. 2009). Far fewer susceptibles survive in this experiment than the previous two. In their 1986 model, Lenski & Hattingh (Lenski & Hattingh 1986) state that there is a set of conditions under which susceptible and resistant strains can coexist only if the resistant bacteria are already established. I suggest that this may be an important factor in previous observations of social resistance, and that rapid and/or prior antibiotic clearance is needed for non-persisters to survive at high antibiotic doses.

3.4.5 Free enzyme facilitates faster antibiotic degradation

The homogenous broth culture environment may facilitate more rapid break down of antibiotic simply by bringing more enzymes into contact with their substrate, as the law of mass action states. In addition, passive extracellular release of β-lactamases from the gram negative periplasm (‘leakage’; Livermore 1995) will have a greater impact in broth culture than in the solid batch culture environment. In my experiments, any exogenous enzyme would diffuse slowly through the agar, reinforcing the spatial structure of the environment, rather

than clearing substantially more antibiotic. Both Clark et al. (2009) and Perlin et al. (2009) find substantially more extracellular β-lactamase in cultures where protective clearance is seen. Curiously, Perlin et al (2009) also suggest that extracellular β-lactamase is more common when *E. coli* is competed with *Salmonella* than when it is grown alone. This could be an indication that *Salmonella* has evolved to exploit the β-lactamase producers, perhaps by causing them to increase production and secretion of β-lactamases, thereby forcing cooperation. Yurtsev et al. (2013) propose that at very high antibiotic doses lysis of resistant cells, leading to the release of β-lactamases into the environment, could speed up antibiotic degradation and increase cooperation. I suggest that if the dose is high enough to lyse resistant cells, susceptible cells will not survive unless there are persisters present.

3.4.6 Protective clearance is not dependent on 'shared' resistance

Dugatkin et al. (2005), Perlin et al. (2009) and Clark et al. (2009) all draw a distinction between 'self-limited' and 'shared' resistance, based on the localisation of the enzyme. The results of this study question this differentiation, as no modifications to pCT have been made to alter the enzyme localisation, which is naturally largely periplasmic (Livermore 1995). Yurtsev et al. (2013) point out that the site of antibiotic action determines the degree to which the resistant cells benefit from the public good, but that this does not affect the overall resistant-susceptible dynamic, which is determined by the rate of degradation alone. More β-lactamase was found in the supernatant of the 'shared' resistance strain, due to breaking off of the enzyme from the cell (Clark et al. 2009), confirming the 'altruistic' nature of resistance compared to the 'self' resistance strain. However, this breaking off from the cell appears to incur a fitness cost for the 'shared' resistance strain. This fact, coupled with free β-lactamase increasing the rate of degradation, could explain the survival of the susceptibles, rather than the enzyme localisation.

3.4.7 Conclusions and outlook

Theory and my data suggest that the conditions for coexistence are quite specific: the initial antibiotic concentration must not exceed the lethal dose, and the rates of degradation and antibiotic influx must not allow the dose to increase (Lenski & Hattingh 1986). Experiments with mercury resistance confirm that frequency dependent coexistence of resistant and susceptible genotypes is strongly dose dependent (Ellis et al. 2007). In my study, using batch culture, no dose greater than the MIC could be found that allowed the coexistence of resistant and susceptible bacteria, unless the susceptibles made the phenotypic switch to persistence.

References

Adam, M. et al., 2008. Epigenetic inheritance based evolution of antibiotic resistance in bacteria. *BMC evolutionary biology*, 8, pp.52–64.

Andrews, J., 2001. Determination of minimum inhibitory concentrations. *Journal of Antimicrobial Chemotherapy*, 48(Suppl S1), pp.5–16.

Balaban, N.Q., 2004. Bacterial Persistence as a Phenotypic Switch. *Science*, 305(5690), pp.1622–1625.

Bigger, J.W., 1944. Treatment of *Staphylococcal* infections with penicillin by intermittent sterilisation. *The Lancet*, 244(6320), pp.497–500.

Brook, I., 2004. β-lactamase-producing bacteria in mixed infections. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, 10(9), pp.777–784.

Brook, I. et al., 1983. In-vivo protection of group A β-haemolytic *Streptococci* from penicillin by β-lactamase-producing *Bacteroides* species. *The Journal of antimicrobial chemotherapy*, 12(6), pp.599–606.

Brooun, A., Liu, S. & Lewis, K., 2000. A dose-response study of antibiotic resistance in *Pseudomonas aeruginosa* biofilms. *Antimicrobial Agents and Chemotherapy*, 44(3), pp.640–646.

Brown, S. et al., 2009. Social evolution in micro-organisms and a Trojan horse approach to medical intervention strategies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364, pp.3157–3168.

Ciofu, O. et al., 2000. Chromosomal β-lactamase is packaged into membrane vesicles and secreted from *Pseudomonas aeruginosa*. *Journal of Antimicrobial Chemotherapy*, 45(1), pp.9–13.

Clark, D., Alton, T. & Bajorek, A., 2009. Evolution of altruists and cheaters in near-isogenic populations of *Escherichia coli*. *Frontiers in Bioscience (Landmark edition)*, 14, pp.4815–4824.

Cottell, J.L. et al., 2011. Complete sequence and molecular epidemiology of IncK epidemic plasmid encoding blaCTX-M-14. *Emerging infectious diseases*, 17(4), pp.645–652.

Cozens, R.M. et al., 1986. Evaluation of the bactericidal activity of β-lactam antibiotics on slowly growing bacteria cultured in the chemostat. *Antimicrobial Agents and Chemotherapy*, 29(5), pp.797–802.

Dahlberg, C. & Chao, L., 2003. Amelioration of the cost of conjugative plasmid carriage in *Eschericha coli* K12. *Genetics*, 165(4), pp.1641–1649.

Dugatkin, L. et al., 2005. Group-beneficial traits, frequency-dependent selection

and genotypic diversity: an antibiotic resistance paradigm. *Proceedings of the Royal Society B: Biological Sciences*, 272, pp.79–83.

Ellis, R. et al., 2007. Frequency-dependent advantages of plasmid carriage by *Pseudomonas* in homogeneous and spatially structured environments. *The ISME journal*, 1(1), pp.92–95.

Gilbert, P., Collier, P.J. & Brown, M.R.W., 1990. Influence of growth rate on susceptibility to antimicrobial agents: biofilms, cell cycle, dormancy, and stringent response. *Antimicrobial Agents and Chemotherapy*, 34(10), pp.1865–1868.

Hackman, A.S. & Wilkins, T.D., 1975. In vivo protection of *Fusobacterium necrophorum* from penicillin by *Bacteroides fragilis*. *Antimicrobial Agents and Chemotherapy*, 7(5), pp.698–703.

Harrison, J.J., Turner, R.J. & Ceri, H., 2005. Persister cells, the biofilm matrix and tolerance to metal cations in biofilm and planktonic *Pseudomonas aeruginosa*. *Environmental Microbiology*, 7(7), pp.981–994.

Hobby, G.L., Meyer, K. & Chaffee, E., 1942. Observations on the mechanism of action of penicillin. *Experimental Biology and Medicine*, 50(2), pp.281–285.

Jones, R.N. & Thornsberry, C., 1982. Cefotaxime: a review of in vitro antimicrobial properties and spectrum of activity. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 4(Supplement 2), pp.S300–S315.

Lenski, R., 1988. Experimental studies of pleiotropy and epistasis in *Escherichia coli*. I. Variation in competitive fitness among mutants resistant to virus T4. *Evolution*, 42(3), pp.425–432.

Lenski, R.E. & Hattingh, S.E., 1986. Coexistence of two competitors on one resource and one inhibitor: a chemostat model based on bacteria and antibiotics. *Journal of theoretical biology*, 122(1), pp.83–93.

Levin, B.R., 1988. Frequency-dependent selection in bacterial populations. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 319(1196), pp.459–472.

Lewis, K., 2007. Persister cells, dormancy and infectious disease. *Nature reviews Microbiology*, 5(1), pp.48–56.

Lewis, K., 2010. Persister Cells. *Annual Review of Microbiology*, 64(1), pp.357–372.

Lewis, K., 2001. Riddle of biofilm resistance. *Antimicrobial Agents and Chemotherapy*, 45(4), pp.999–1007.

Livermore, D.M., 1995. β-Lactamases in laboratory and clinical resistance. *Clinical Microbiology Reviews*, 8(4), pp.557–584.

McKenney, D., 1997. Influence of growth rate and nutrient limitation on susceptibility of *Burkholderia cepacia* to ciprofloxacin and tobramycin. *Journal of Antimicrobial Chemotherapy*, 40(3), pp.415–417.

Neu, H.C., 1982. The in vitro activity, human pharmacology, and clinical effectiveness of new β-lactam antibiotics. *Annual review of pharmacology and toxicology*, 22(1), pp.599–642.

Perlin, M.H. et al., 2009. Protection of *Salmonella* by ampicillin-resistant *Escherichia coli* in the presence of otherwise lethal drug concentrations. *Proceedings of the Royal Society B: Biological Sciences*, 276(1674), pp.3759–3768.

R Core Team, 2013. R: A language and environment for statistical computing. Available at: http://www.R-project.org.

Rolinson, G., Macdonald, A. & Wilson, D., 1977. Bactericidal action of β-lactam antibiotics on *Escherichia coli* with particular reference to ampicillin and amoxycillin. *Journal of Antimicrobial Chemotherapy*, 3(6), pp.541–553.

Tacking, R., 1954. Penicillinase-producing bacteria in mixed infections in rabbits treated with penicillin. 2. Studies on penicillinase-producing *coli*- and *pyocyaneus*-bacteria. *Acta pathologica et microbiologica Scandinavica*, 35(5), pp.445–454.

Tuomanen, E. et al., 1986. The rate of killing of *Escherichia coli* by β-lactam antibiotics is strictly proportional to the rate of bacterial growth. *Microbiology*, 132(5), pp.1297–1304.

West, S.A. et al., 2006. Social evolution theory for microorganisms. *Nature reviews Microbiology*, 4(8), pp.597–607.

West, S.A. et al., 2007. The social lives of microbes. *Annual review of ecology, evolution, and systematics*, 38(1), pp.53–77.

Yurtsev, E.A. et al., 2013. Bacterial cheating drives the population dynamics of cooperative antibiotic resistance plasmids. *Molecular Systems Biology*, 9, pp.1–7.

Chapter 4: Plasmid sequencing

4.1 Introduction

4.1.1 The 'cheat' plasmid hypothesis

Antibiotic resistance conferred by the enzymatic breakdown of drugs, such as β-lactamases, has the potential to be a cooperative trait (Chapter 3, Section 3.1.1). These antibiotic resistance genes are often carried on plasmids. Horizontal transfer of plasmid carried resistance genes to plasmid-free cells would ensure that all cells share the cost of public good (β-lactamase) production, and prevent the evolution of social cheats: cells which benefit from the enzyme produced by co-operators, but which do not themselves produce the public good. Plasmid incompatibility groups and mating types can affect the spread of plasmids, leading to the idea of 'cheat' plasmids (Figure 4.1). A 'cheat' plasmid is of similar size to the resistance plasmid, carrying similar genes and in the same incompatibility group, but crucially it does not carry the resistance gene. Carriage of the 'cheat' plasmid would prevent cells from obtaining the resistance plasmid via conjugation, facilitating the existence of susceptible cheats in the population, which can exploit the public good produced by other cells. A cheat-type plasmid has been successfully competed with resistance-plasmid carrying cells, by replacing the plasmid's ampicillin resistance gene with kanamycin resistance (Dugatkin et al. 2005).

It has been proposed that 'Trojan horse' susceptible cheats could invade and displace resistant populations, facilitating clearance of an infection with antibiotics (Brown et al. 2009). 'Cheat' plasmids could block the spread of resistance genes by horizontal transfer, and thus invade resistant populations in this way. In addition, incompatibility-based plasmid displacement using artificial plasmid constructs carrying key features of the displacement target has been successful for plasmid curing (Hale et al. 2010), confirming that incompatibility with highly similar plasmids is a viable mechanism for manipulation of plasmid carriage. Potential 'cheat' plasmids have been identified in environmental and

veterinary isolates by collaborators at the Animal Health and Veterinary Laboratories Agency (AHVLA).

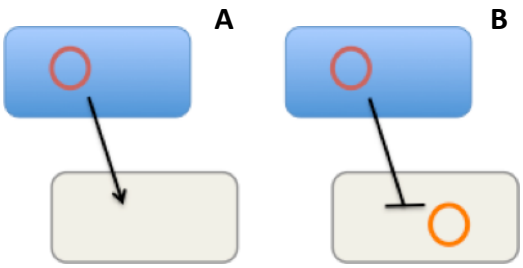

Figure 4.1: 'Cheat' plasmids prevent the transfer of resistance plasmids to susceptible cells. (A) Resistant cells (blue) can transfer resistance plasmids to plasmid-free susceptible cells (white). **(B)** Susceptible cells bearing the 'cheat' plasmid cannot receive the resistance plasmid.

4.1.2 Plasmids used in this chapter

The extended-spectrum β-lactamase resistance plasmid, pCT, was identified from a veterinary outbreak of *E. coli* C159/11, by Liebana *et al*. (2006). Four additional *E. coli* clones were isolated from environmental (strain 24456) and veterinary (strains Cow63, Cow52 and Cow23) samples from the same outbreak farm. These strains appear to carry plasmids of similar size to pCT, and were confirmed as positive using a pCT specific PCR screening (M. O. Stokes, personal communication, (Stokes et al. 2012)). Little is known about these plasmids, other than that they do not carry the CTX-M-14 resistance gene present in pCT. Their apparent similarity to pCT, coupled with the absence of the resistance gene, made them ideal candidate 'cheat' plasmids.

4.1.3 Aims

The aims of this chapter were firstly, to sequence the four plasmids isolated from the same outbreak site as pCT: p24456, pCow63, pCow52 and pCow23; and secondly, to annotate and characterise these plasmids, to determine their potential 'cheat' status and to identify any genes or features of interest. In order

to be a potential cheat, a plasmid must be of the same incompatibility group as pCT (IncK), and must not carry the resistance gene CTX-M-14.

4.2 Methods

4.2.1 Sample preparation and sequencing

Plasmids were isolated from the four strains using a High Speed Midi Kit (Qiagen), and an agarose gel of the extractions was run as described by Gonzalez et al. (1982). Plasmid extractions were stored in TE buffer at -20°C until sequencing. Extracted plasmid DNA was quantified fluorometrically using Quant-iT PicoGreen dsDNA assay kit (Invitrogen, Life Technologies Ltd., Paisley, UK) and POLARstar Galaxy fluorometer (BMG Labtech Ltd., Aylesbury, UK) and diluted to approximately 5 ng μl^{-1} DNA in TE buffer. Sequencing was conducted using the 454 GS-FLX Titanium series sequencer (Roche Diagnostics Ltd., Basel, Switzerland). Libraries were prepared from the plasmid DNA using the Rapid Library Preparation Kit, as specified by the Rapid Library Preparation Method Manual (Roche Diagnostics Ltd., October 2009). Briefly, DNA was fragmented by nebulisation, and the four libraries tagged with unique multiplex identifier adaptors, allowing them to be processed as one pooled library. This library was analysed for average fragment size and concentration using the high sensitivity DNA kit and 2100 Bioanalyser (Agilent Technologies Ltd., Santa Clara, USA). DNA concentration was optimised for emPCR (emulsion PCR), to obtain optimum bead enrichment of around 8%, and finally emPCR was conducted to produce enriched beads for sequencing. These enriched beads were then taken forward for sequencing, as described in the Roche GS FLX Titanium Series Sequencing Method Manual (Roche Diagnostics Ltd., October 2009). Initial assembly of sequences was done using the GS de novo Assembler (Newbler) program (Roche Diagnostics Ltd).

4.2.2 Contig closing

Initial assembly produced contigs of varying size for the four plasmids. PCR was used to join contigs together and circularise the sequences of p24456 and pCow63. Primers were designed using existing contig sequences, with primers directed outwards from the contig ends, testing all possible contig joins. Contig orientations were altered using Geneious (Biomatters 2012), and primers designed using Primer3 (Rozen & Skaletsky 2000). Table 4.1 gives primer

sequences and individual melting temperatures (T_m). Annealing temperatures were determined for each primer pair combination by subtracting 5°C from the mean T_m. For full PCR protocol and conditions see Chapter 2, Section 2.2.1. Primer combinations that produced a product were then sequenced. The PCR products were cleaned up using Exo1 (Fisher Scientific) and TSAP (Promega UK) both at final concentration 0.045 units μl^{-1}. Clean up mixes were incubated at 37°C for 30 minutes, followed by 80°C for 15 minutes. Sequencing reactions were conducted using BigDye® terminator v3.1 cycle sequencing kit (Life Technologies Ltd.) with both forward and reverse diagnostic primers. Cycling conditions were as follows: 96°C for 1minute, followed by 25 cycles of: 96°C for 10 seconds, 50°C for 5 seconds, 60°C or 4 minutes. Sequencing reactions were cleaned up using ethanol/sodium acetate precipitation. The sequencing runs were conducted by AHVLA Scientific (AHVLA, Weybridge, UK).

Plasmid	Name	Sequence (5'-3')	T_m (°C)
p24456	contig 1 F	CCCAGCAGATCTGAGAGTCC	61.4
	contig 1 R	AAATGATTTTCACCGGCAAG	53.2
	contig 2 F	ACGGATGCACCTGAACTACC	59.4
	contig 2 R	GCAAGTTCAGAGGCAGGAAC	59.4
pCow63	contig 1 F	AGCGCAGGATGTGAGCTAAT	57.3
	contig 1 R	GTACGGGCTGCTATCTCTGC	61.4
	contig 2 F	CGTAATTACCACCGGCAGAT	57.3
	contig 2 R	GATTGAATGCCCCTTGAGAA	55.3
	contig 3 F	CCTGGATGCGACAGGTAAAT	57.3
	contig 3 R	GCCTCTGTAAGCCTTTTTCG	57.3
	contig 4 F	CAGACCAGTTGGCAGAATCA	57.3
	contig 4 R	GGTGAAAGATCCCCCTGTT	57.3

Table 4.1: Contig closing primer sequences with melting temperatures.

4.2.3 Sequence analysis

All contigs larger than 1000 bp were screened using NCBI Nucleotide BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to exclude chromosomal contamination. Contigs were excluded only if BLAST matches were clearly chromosomal

contaminants i.e. top 10 matches having high identity (>95%) to chromosomal loci, and no plasmid matches present. Matches to phage or transposon genes were retained at this stage. Contigs smaller than 1000 bp were excluded. Alignment of contigs and contig-closing sequences were carried out using Geneious (Biomatters 2012).

Sequence and structural homology between the partially closed plasmid sequences and pCT was assessed using Mauve (Darling et al. 2010). Annotation of the sequences was carried out using RAST (Overbeek et al. 2014) and Artemis (Rutherford et al. 2000; Berriman & Rutherford 2003), and sequences were visualised using DNAPlotter (Carver et al. 2009). In addition, the sequence of pCow63 was uploaded to the Joint Genome Institute's (JGI) Integrated Microbial Genomes (IMG) database (Markowitz et al. 2014), which provides annotation and comparison tools, facilitating comparison with genomes and metagenomes held in the database. This also allowed comparison between the IMG and RAST annotations of pCow63.

4.3 Results

4.3.1 Plasmid identification

Plasmids of appropriate size (~93 Kb) were successfully isolated from the four *E. coli* strains (Figure 4.2). The gel shows two bands of high M_w (above the 10 Kb marker) for all strains, probably due to the supercoiled and open circular plasmid formations. Extractions from Cow63 and Cow23 have additional lower M_w bands at around 4 Kb and 1 Kb respectively, due either to the presence of additional small plasmids in the extraction, or to contamination with small genomic DNA fragments. Genomic and small plasmid contaminants were excluded after sequencing.

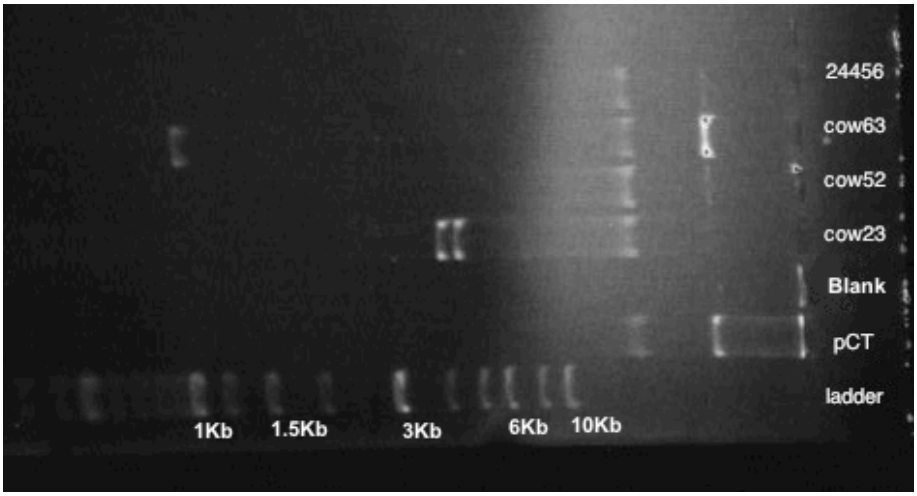

Figure 4.2: Gel showing plasmid extractions from *E. coli* strains. pCT is included for comparison, and is approximately 93 Kb. Lane 1 contains 2-log ladder (NEB).

4.3.2 pCow52/24456 assembly

Sequence assembly with the GS de novo assembler (Roche Diagnostics Ltd.) produced one large contig of 103643 bp for pCow52 and two contigs of 85629 bp and 18013 bp for p24456 (total length 103642 bp). All but one of the smaller contigs (>1000 bp) were identified as genomic contamination using BLAST and were therefore excluded. p24456 had one 1250 bp contig identified as a transposase/prophage DNA. The Mauve alignment indicates very high homology

between pCow52 and p24456 (Figure 4.4, rows 1 and 2), with the exception of the smallest, 1250 bp, contig of p24456. Subsequent Geneious alignment showed just four SNPs between p24456 and pCow52. Three of these were additional nucleotides in single-base runs, located in the pCow52 plasmid at bases 6310 (A), 92630 (T) and 99949 (T). These single base runs are likely to be error prone both within the cell and in the sequencing run. There is just one transition substitution; an A in pCow52 at position base 8639 replaces a G in p24456.

Contig closing PCR and sequencing successfully joined contig 1 (85629 bp) and contig 2 (18013 bp) of p24456, with primers *contig 1 F* and *contig 2 R* (Table 4.2), producing a 432 bp sequencing product which perfectly aligned to the contigs 1 and 2, creating a linear plasmid (Figure 4.3). Although primers *contig 1 R* and *contig 2 F* did produce a PCR product, sequencing from the forward and reverse primers did not overlap, so it is not possible to confirm this join. Unfortunately the alignment with pCow52 also fails to confirm this join, as there are just four bases of potential overlap between the contig ends. Further PCR and sequencing is required to fully circularise this plasmid.

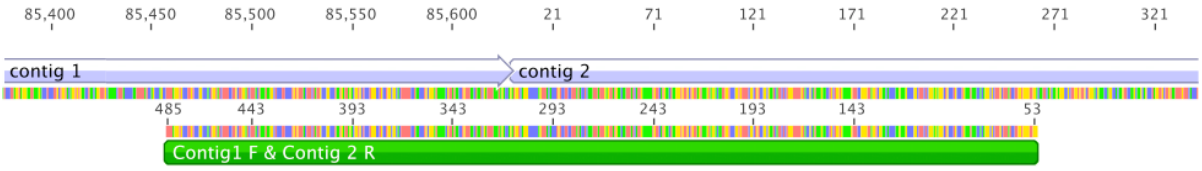

Figure 4.3: Contig closing of p24456. This figure shows the alignment of concatenated contigs 1 & 2 (top), with the 432 bp consensus sequence produced by primers contig 1 F and contig 2 R (bottom).

The fact that PCR products were generated with both primer sets, joining contigs 1 and 2, and that BLAST identification suggested phage or transposase DNA indicates that contig 3 (1250 bp) is not part of this plasmid, and is probably genomic, small plasmid or phage contamination. From this we can conclude that these plasmids are the same, and henceforth will be referred to as pCow52/24456.

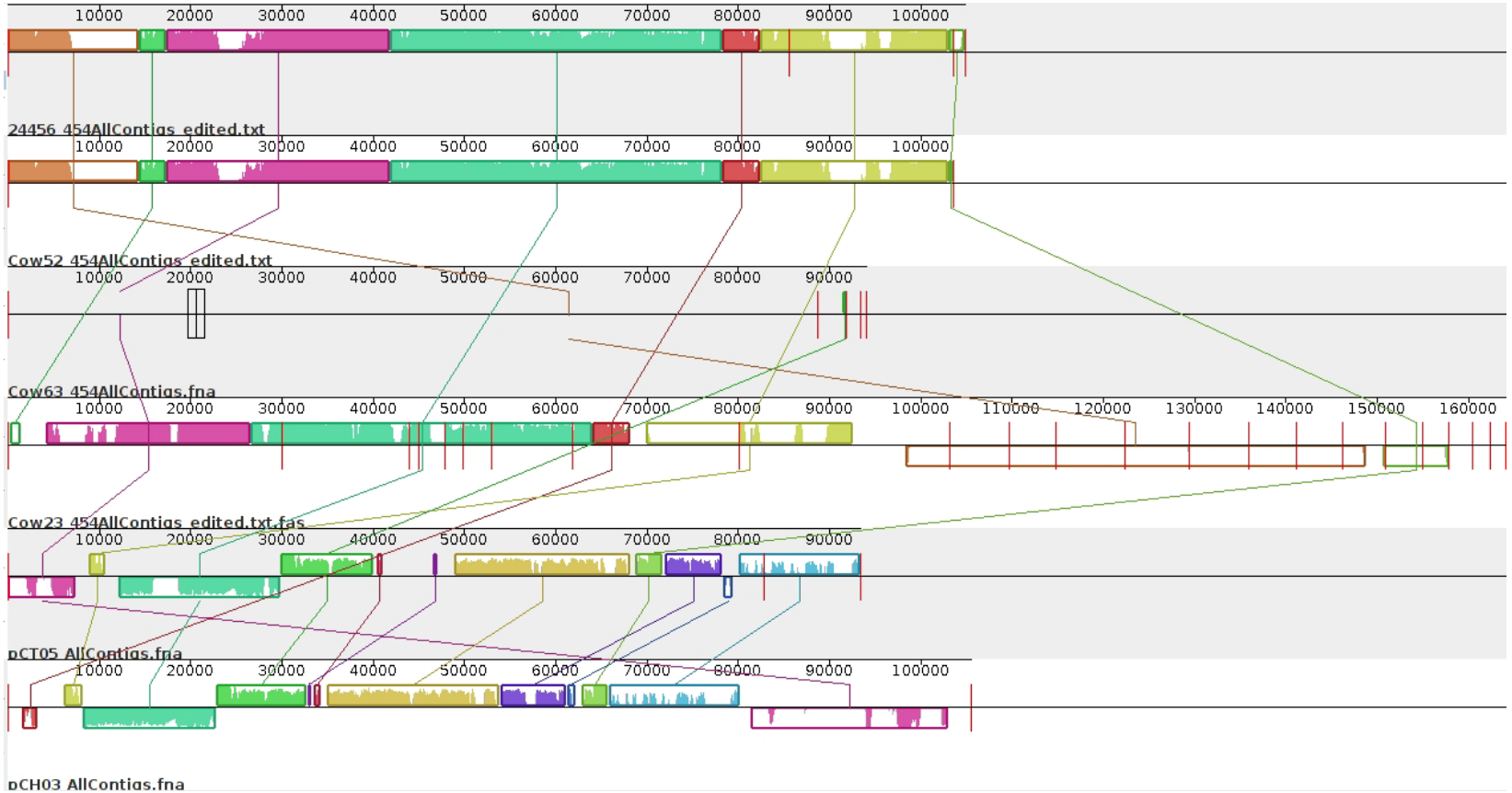

Figure 4.4: Mauve alignment of plasmid sequences. Each horizontal line represents a plasmid sequence, from top to bottom these are: p24456, pCow52, pCow63, pCow23, pCT05 and pCH03. pCT05 is the sequence of the resistance plasmid pCT, and pCH03 is an unrelated plasmid sequence for comparison (from a chicken isolate carrying the β-lactamase CTX-M-1; not linked to the outbreak farm in this study, R. J. Ellis, personal communication). Coloured blocks indicate regions of alignment between sequences, the colour intensity within the blocks represents the similarity profile; the average level of conservation in that region of the genome sequence. Connector lines join corresponding blocks in different sequences. Blocks above the centre line align in the forward direction, and blocks below the line align in the reverse orientation. Red vertical lines separate contigs.

4.3.3 Features of pCow52/24456

RAST identified 128 coding sequences (CDS) in pCow52/24456, whereas Artemis detected 228 CDS. This disparity is largely explained by 102 CDS categorised as 'no product' by Artemis. Figure 4.5 shows all CDS identified by Artemis, with some key features marked. The following sections highlight some key plasmid maintenance and virulence genes identified on pCow52/24456 and the other two plasmids, but these are not exhaustive lists. For a full list of CDS identified by RAST for all plasmids, please see Appendix 1.

Surprisingly, pCow52/24456 is not an IncK incompatibility group plasmid, as indicated by previous work (M. O. Stokes, personal communication). It is in fact an IncF plasmid, carrying replication proteins *repA1* and *repA3* of FII replicon and *repE* of FIA replicon (Villa et al. 2010). pCow52/24456 carries 28 plasmid transfer genes (22 *tra* and 6 *trb* genes), with almost identical structure to the standard IncF transfer region (Willetts & Skurray 1980; Lawley et al. 2004). Artemis, but not RAST, detected the *traS* gene and curiously neither detected *trbA*. BLAST searching of the region where *trbA* should be (between traQ and traF) also failed to identify any genes. *TrbA* is of unknown function, and insertion mutations within this gene do not significantly affect transfer ability, suggesting that pCow52/24456 is likely to be conjugative (Frost et al. 1994; Lawley et al. 2004). Several other archetypal plasmid features were identified including: the classical toxin/antitoxin plasmid addiction system *ccdA*/*ccdB* (Kobayashi 2004); the plasmid partitioning proteins *parA*/*parB* (Funnell & Slavcev 2004); and a common feature of IncF plasmids, the non-essential but highly conserved *ssb,* single stranded binding protein, and *psiA/B* SOS stress response inhibitor (Lawley et al. 2004). Neither annotation program identified the partitioning related site *parS* which should be located downstream of *parA/parB* and is essential for plasmid stability. RAST also identified an ABC-type bacteriocin transporter protein, however sequence searching with the BAGEL3 bacteriocin mining software failed to identify any bacteriocin genes (van Heel et al. 2013). In addition to these plasmid maintenance genes, pCow52/24456 also carries the iron scavenging operon, *iroA*, at position 8589-18156, which encodes genes for salmochelin siderophore production and uptake (Lin et al. 2005).

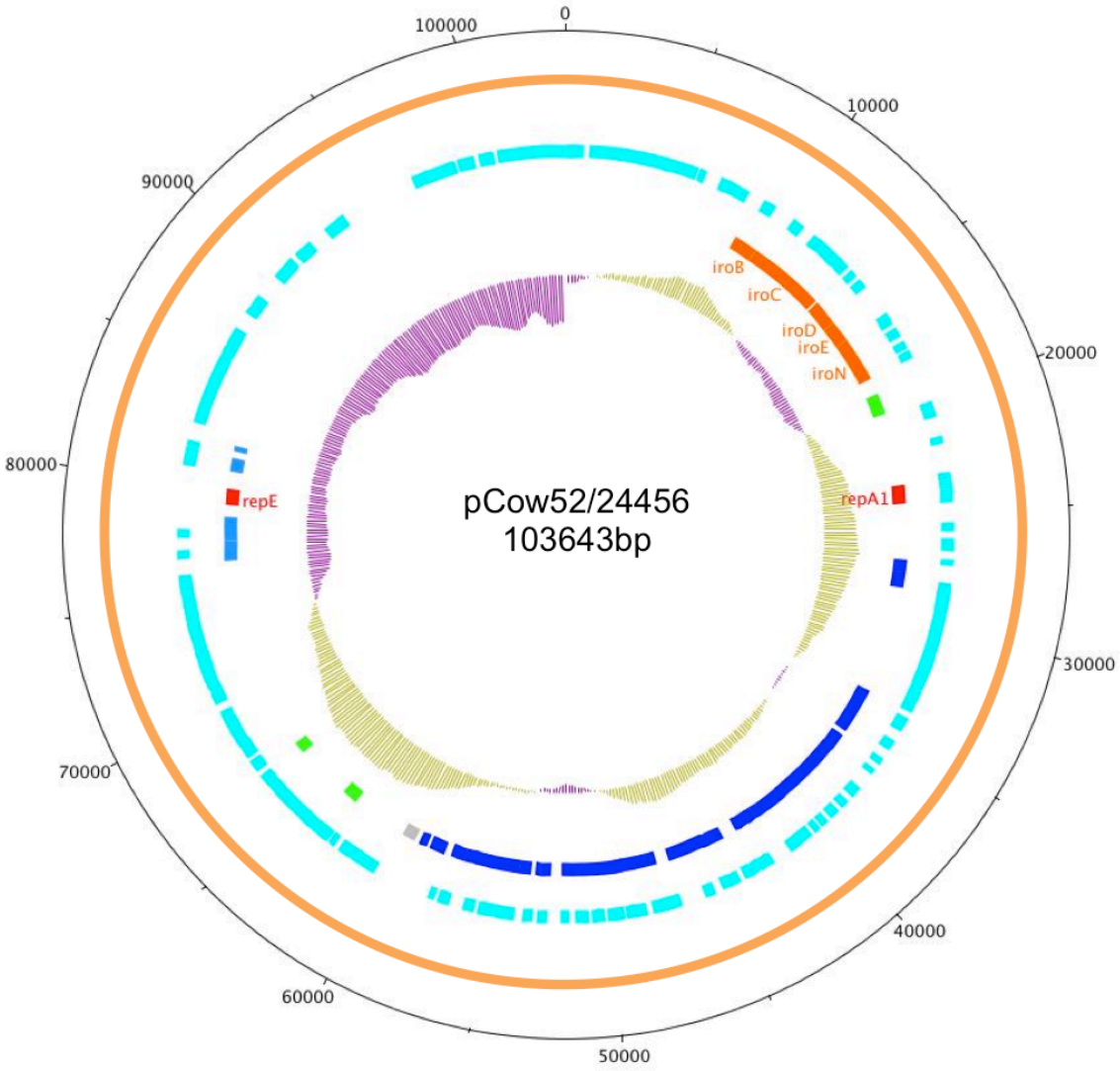

Figure 4.5: pCow52/24456. All coding sequences (CDS) detected by Artemis are shown in turquoise. The inner row highlights some genes of interest. Dark blue: plasmid transfer; Light blue: plasmid maintenance/replication; Orange: virulence; Red: Rep genes; Green: miscellaneous other. The graphs indicate sequence GC ratio: purple < average, yellow > average.

4.3.4 Assembly and features of pCow23

Sequencing of pCow23 produced 22 contigs of plasmid origin >1000 bp, with four >10000 bp, the largest of these being 30121 bp, and a total length of >

160000 bp. Contig closing was not attempted for pCow23 due to the large number of contigs, and it was therefore analysed as a concatenated sequence. The Mauve alignment shows large regions of homology between pCow23 and pCow52/24456, particularly in the red, fuschia and teal coloured boxes, which also show some homology to pCT (Figure 4.4, row 4). This alignment also shows 13 contigs with no homology to the other plasmids. Although BLAST searches identified all of these contigs as being of plasmid origin, the initial plasmid extraction showed two bands of approximately 4 Kb. It is possible that some of these contigs represent smaller plasmid contamination. However, the unique sequence represents approximately 60 Kb, and it seems unlikely that this would be entirely the result of contamination with a 4 Kb plasmid.

As with pCow52/24456, RAST annotation identified fewer CDS than Artemis annotation, just 211 compared to 366. However, 260 of these were designated 'no product', and only 2 potential gene products were identified outside the four largest contigs (1-4, respectively 30121, 23107, 18219 and 13852 bp). Like pCow52/24456, pCow23 is an IncF plasmid, carrying *repA1* of the FII replicon *a*nd *repE* of FIA replicon*,* but unlike pCow52/24456 it also carries *repA* of FIB replicon (Villa et al. 2010). pCow23 carries an almost complete IncF transfer region, split across contig 1, *traX- traN*, and contig 4*, traC- traM*. There is just one, non-essential, gene missing; trbG, and there are two copies of both *traM* and *traC* (Lawley et al. 2004). The division of the transfer region across contigs 1 and 4 suggests that perhaps these contigs should be joined.

There are several other similarities between pCow23 and pCow52/24456: IncF features such as *ssb* and *psiA/B*; the stability locus of IncFII plasmid NR1 and RepA2 *copB* regulatory protein (Tabuchi et al. 1992; Tabuchi et al. 1988; Dong et al. 2004); and the toxin/antitoxin system *ccdA*/*ccdB.* pCow23 also appears to have a second post-segregational killing system, *mvpT*/*mvpA*, located within the *trbH* region, as described by Sayeed et al. (2000). Other plasmid-associated genes were also identified, including the DNA binding protein, *artA,* which regulates horizontal and vertical plasmid transmission (Ni et al. 2009), and the error-prone DNA polymerase V, *umuC*, (the plasmid version is called *mucB* (Goldsmith

et al. 2000)) which is involved in stress response (P. Reuven & Eldar 2011; N. B. Reuven et al. 1999; Maor-Shoshani et al. 2000). In addition, RAST annotation identified a colicin D gene on contig 7 (position 738- 2381), which was confirmed using BAGEL3 (van Heel et al. 2013) (Figure 4.6). Although the BAGEL3 annotation calls ORF 008 'Pyocin S', BLAST searching confirms that it is in fact colicin D (Roos et al. 1989). The S-type pyocins of *Pseudomodas pyogenes* are highlight similar to colicins in their domain structure, so this is the likely cause of confusion in annotation (Cascales et al. 2007). ORF 003 of contig 7 encodes the essential colicin immunity gene.

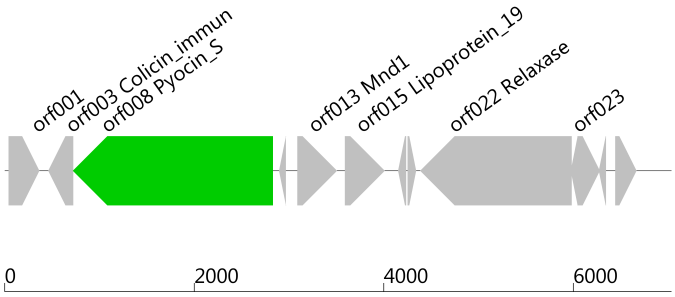

Figure 4.6: pCow23 contig 7 carries colicin B. ORF 008 encodes colicin B, and ORF 003 the corresponding colicin immunity gene. Image produced using BAGEL3 (van Heel et al. 2013).

Alongside these plasmid maintenance genes, pCow23 also carries a number of virulence-associated genes: the putative fluoride ion transporter/camphor resistance gene *crcB* (Sand 2003); the *Salmonella* virulence genes *vagC* and *vagD* (Pullinger & Lax 1992); the manganese uptake operon *sitA, -B, -C, -D* (Davies & Walker 2007; Sabri 2006); the aerobactin siderophore synthase genes, *iucA/D*, and a partial siderophore receptor operon *iutA* (de Lorenzo et al. 1986); the K88 fimbrial genes *faef, -g, -h, -i, -j*, associated with binding to the pig intestine (Bakker et al. 1992); and finally *espC*, an enterotoxin associated with EPEC (enteropathogenic *E. coli*) strains (Salinger et al. 2008).

Multiple gene copies are present on pCow23. This repetition is likely due to the large number of contigs that have not been successfully closed. Re-sequencing of

the plasmid would allow contigs to be assembled more completely, and would identify true gene duplications.

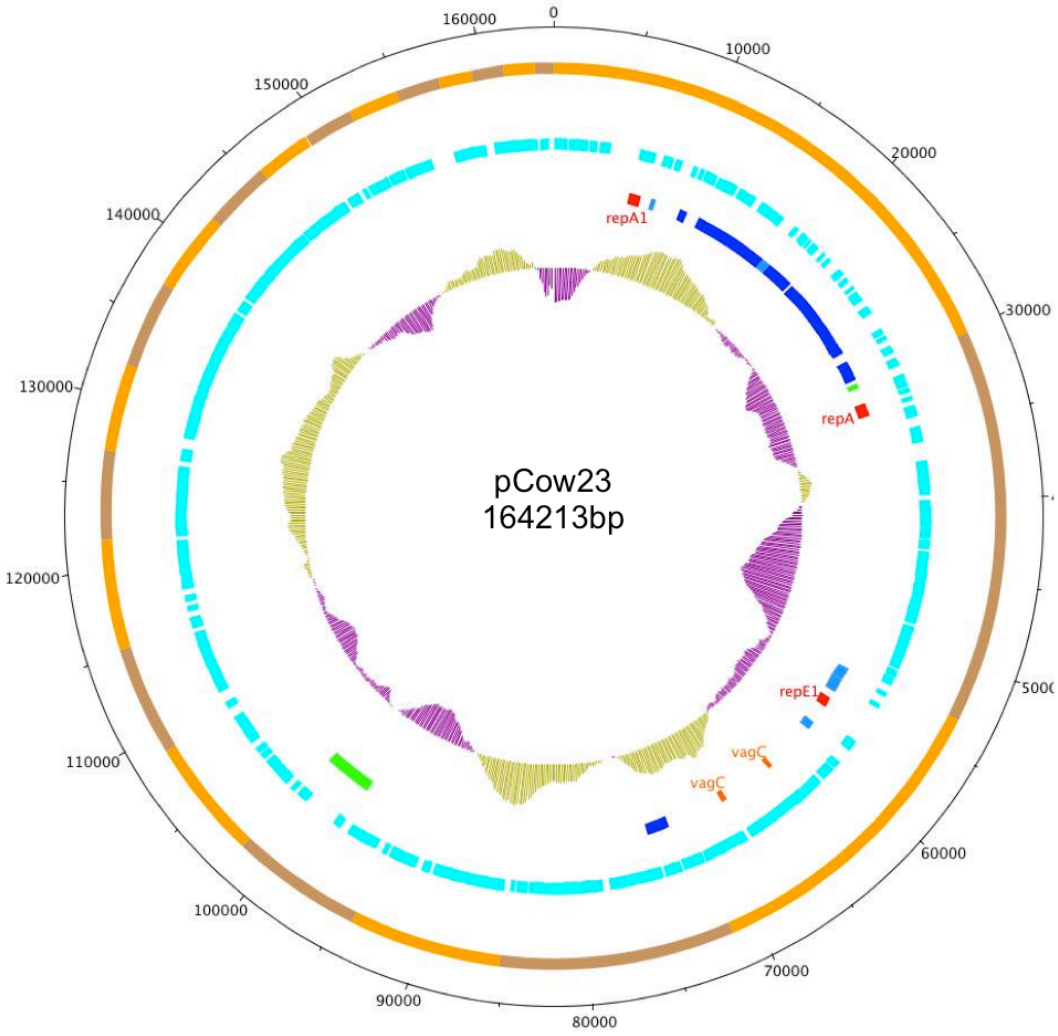

Figure 4.7: pCow23. Orange/brown line indicates individual contigs. All coding sequences (CDS) detected by Artemis are shown in turquoise. The inner row highlights some genes of interest. Dark blue: plasmid transfer; Light blue: plasmid maintenance/replication; Orange: virulence; Red: Rep genes; Green: miscellaneous other. The graphs indicate sequence GC ratio: purple < average, yellow > average.

4.3.5 Assembly and features of pCow63

Sequencing of pCow63 produced four contigs of 88710, 3176, 1551 and 670 bp. Contig joining PCR and sequencing connected contigs 2, 4 and 1 in that order. Successful primer pairs and their resulting product sizes are given in Table 4.2. Sequencing of the product from *contig 1 F* and *contig 2 R* primers was inconclusive, as sequencing from the forward and reverse primers did not overlap. However, primers designed for contig 3 (Table 4.1) formed no product with any other primer. It therefore seems likely that contigs 1 and 2 are contiguous. This could be confirmed with further PCR and sequencing. Given these results, the 1551 bp contig 3 was excluded from further analysis of pCow63, giving a total length 92556 bp. The plasmid extraction (Figure 4.2) shows a band of approximately 1 Kb. It is possible that this corresponds to contig 3, and may be a smaller plasmid present in the extraction. RAST annotation of this contig identified just three CDS; two hypothetical proteins and the plasmid replication associated gene, *repA*, indicating that this is indeed plasmid DNA.

Primer 1	**Primer 2**	**Product (bp)**
Contig 2 F	Contig 4 R	677
Contig 4 F	Contig 1 R	752
Contig 1 F	Contig 2 R	1300

Table 4.2: Contig joining primers for pCow63.

The uniqueness of pCow63 can clearly be seen in the Mauve alignment (Figure 4.4, row 3). It shows almost no homology with the other plasmids in this study, with pCT05 or with pCH03. BLAST searches reveal that it also has very little homology with other sequenced plasmids. pCow63 was annotated using RAST and also the JGI IMG/ER annotation pipeline. The outcomes of the two tally well, with RAST identifying 117 CDS (contig 4 was omitted in error), and IMG/ER identifying 122 (Figure 4.8). Just 18 of these were genes encoding known proteins or functions, 33 CDS were of phage or pro-phage origin and the remainder were hypothetical proteins and genes with unknown functions. Very few common plasmid features were identified: the *repE* gene of the FIA replicon,

and the FIC replicon associated *repL* gene, the plasmid stability genes *parA & B* and two separate toxin/antitoxin loci: *phd/doc* and *higA/B* (Figure 4.8). The *phd/doc* addiction system is associated with prophage P1, as are *clpX* and *xerD*. *XerD* is a phage specific recombinase and *clpX* is ATP-dependent protease, involved in degradation of the *doc* toxin (Lehnherr & Yarmolinsky 1995; Lehnherr et al. 1993). pCow63 also carries genes that are not commonly associated with either plasmids or phage including genes for glutamate metabolism, nucleotide metabolism (dUTPase), a stomatin/prohibitin-like membrane protease, and an N6 adenine-specific DNA methylase.

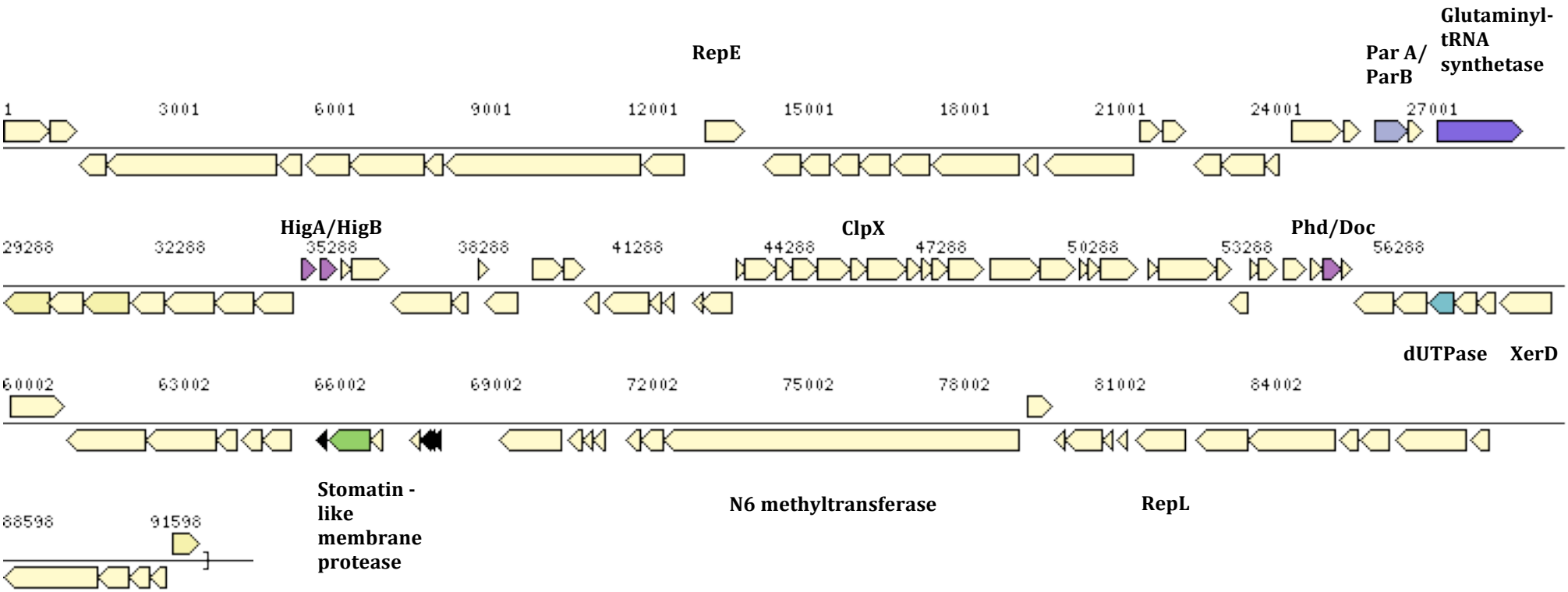

Figure 4.8: pCow63. Contigs are shown in the order 2, 4, 1 with a total length 92556 bp. Cream colour indicates genes involved in replication/recombination & repair, but the majority of these are unknown or hypothetical proteins. Some key genes identified by both RAST and IMG/ER are highlighted in colour and gene names/products given. Black indicates RNA. Figure produced using JGI IMG/ER chromosome viewer (Markowitz et al. 2014).

4.4 Discussion

4.4.1 These are not potential 'cheat' plasmids

Perhaps the most pertinent fact uncovered through sequencing of these four plasmids is that they are not in the IncK incompatibility group, and are therefore not candidate 'cheat' plasmids for competition with pCT. Both pCow52/24456 and pCow23 are IncF plasmids. It is likely that pCow63 is also IncF, as it carries *repE* and *repL* genes associated with IncF, but it is difficult to draw firm conclusions about pCow23 with so few plasmid-associated genes to work with. With hindsight, it would have been wise to ascertain the plasmids' incompatibility before sequencing, perhaps using the replicon typing PCR method of Carattoli et al. (2005) used in Chapter 5 of this thesis. However, some interesting and unusual features were found in these plasmids, which are discussed below.

4.4.2 IncF plasmids and virulence

All the plasmids sequenced here are in the IncF incompatibililty group. IncF plasmids are generally associated with *Enterobacteriaceae,* and are one of the most common plasmid types, found across a range of *E. coli* hosts from avian pathogenic strains to human commensal strains (Johnson et al. 2007). They are usually large, >100 Kb, often carry multiple replicons, and have been associated with virulence genes such siderophores and the CTX-M-15 β-lactamase (Villa et al. 2010). pCow52/24456 and pCow23 fit this image of IncF plasmids well; both are large and carry multiple replicons and virulence genes. The *iroA* operon of pCow52/24456 has been shown to be an important virulence factor in avian pathogenic *E. coli* strains, and has been found in avian CTX-M carrying *E. coli* (Stokes et al. 2012; Caza et al. 2008). pCow23 carries a multitude of virulence factors, including siderophore and manganese uptake operons, as well as EPEC associated features such as fimbriae and the *espC* entertoxin (Salinger et al. 2008). K88 fimbriae have been associated with binding to the pig intestine, which is surprising given that this is a plasmid isolated from a cattle farm (Bakker et al. 1992). However, further information about other livestock on the

farm is not available, so it is not possible to determine whether this plasmid may have transferred from another host animal, or whether these fimbriae are present in bovine enteric *E. coli*. Phenotype or gene-expression assays are required to determine which (if any) of the virulence genes identified are functional, especially for those genes that appear to be in partial operons, such as the pCow23 aerobactin uptake operon. Further information about the plasmids' host strains would also be useful for determining the functionality of plasmid genes. However, these strains were isolated from the site of a veterinary *E. coli* outbreak (Liebana et al. 2006), so it seems likely that there may be other virulent strains present.

4.4.3 pCow63 is unique

pCow63 appears to be a unique plasmid. It shows little similarity to the other plasmids in this study, and carries few genes of known function. It appears to be largely comprised of phage-associated DNA, with very few plasmid maintenance genes and no discernable host beneficial genes. The functional genes it does carry appear to be a mishmash of bacterial metabolic genes that are not generally carried on plasmids. This could be a result of recombination with the host, or it is possible that these genes were transposed via faulty phage replication and insertion. Although pCow63 is not a candidate 'cheat' plasmid for pCT, as it is not IncK, it could potentially be a cheat plasmid in another system, as it does not carry any virulence or resistance genes. It is also an example of a potential cryptic plasmid: a plasmid that confers no apparent benefit to its host. Curing the host strain of pCow63 and comparing any fitness changes could confirm this. Sequencing of this plasmid highlights just how little we still know about plasmids: what is its function (if any)? How is it maintained? Why has it not been lost? These themes are discussed further in Chapter 6 of this thesis.

4.4.4 Conclusions and outlook

This study has identified two IncF plasmids, pCow52/24456 and pCow23, and a novel plasmid, pCow63, from the same sample site as the IncK CTX-M-14 carrying plasmid, pCT. The variation found in this study demonstrates the

diversity of plasmids being sustained by the *E. coli* population at the isolation farm. The fact that both resistance and non-resistance plasmids, and multiple plasmids of the same incompatibility group are maintained is an interesting observation about the ability of bacterial populations to sustain plasmid 'parasites'. It is also interesting to note that pCow52/24456 is found in both veterinary (Cow52) and environmental (24456) isolates, indicating that plasmids may pass between these distinct communities. Further contig joining and re-sequencing is required to complete the sequencing of these plasmids, but as draft sequences they can still provide insights into the plasmid diversity of the outbreak farm. Although not possible within the scope of this thesis, curing of pCow63 carrying strains coupled with gene knockouts and phenotypic analysis could yield interesting results about the unknown features of this plasmid.

References

Bakker, D. et al., 1992. Identification of minor fimbrial subunits involved in biosynthesis of K88 fimbriae. *Journal of Bacteriology*, 174(20), pp.6350–6358.

Berriman, M. & Rutherford, K., 2003. Viewing and annotating sequence data with Artemis. *Briefings in Bioinformatics*, 4(2), pp.124–132.

Biomatters, 2012. Geneious version 5.2. Available at: http://www.geneious.com/.

Brown, S. et al., 2009. Social evolution in micro-organisms and a Trojan horse approach to medical intervention strategies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364, pp.3157–3168.

Carattoli, A. et al., 2005. Identification of plasmids by PCR-based replicon typing. *Journal of microbiological methods*, 63(3), pp.219–228.

Carver, T. et al., 2009. DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics (Oxford, England)*, 25(1), pp.119–120.

Cascales, E. et al., 2007. Colicin biology. *Microbiology and molecular biology reviews*, 71(1), pp.158–229.

Caza, M. et al., 2008. Specific roles of the *iroBCDEN* genes in virulence of an avian pathogenic *Escherichia coli* O78 strain and in production of salmochelins. *Infection and immunity*, 76(8), pp.3539–3549.

Darling, A.E., Mau, B. & Perna, N.T., 2010. progressiveMauve: Multiple genome alignment with gene gain, loss and rearrangement. *PloS one*, 5(6), p.e11147.

Davies, B.W. & Walker, G.C., 2007. Disruption of *sitA* compromises *Sinorhizobium meliloti* for manganese uptake required for protection against oxidative stress. *Journal of Bacteriology*, 189(5), pp.2101–2109.

de Lorenzo, V. et al., 1986. Aerobactin biosynthesis and transport genes of plasmid ColV-K30 in *Escherichia coli* K-12. *Journal of Bacteriology*, 165(2), pp.570–578.

Dong, X. et al., 2004. PlasMapper: a web server for drawing and auto-annotating plasmid maps. *Nucleic Acids Research*, 32(Web Server issue), pp.W660–W664.

Dugatkin, L. et al., 2005. Group-beneficial traits, frequency-dependent selection and genotypic diversity: an antibiotic resistance paradigm. *Proceedings of the Royal Society B: Biological Sciences*, 272, pp.79–83.

Frost, L.S., Ippen-Ihler, K. & Skurray, R.A., 1994. Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiological reviews*, 58(2), pp.162–210.

Funnell, B.E. & Slavcev, R.A., 2004. Partitioning systems of bacterial plasmids. In B. E. Funnell & G. J. Phillips, eds. *Plasmid Biology*. Washington D.C.: American Society for Microbiology (ASM), pp. 81–103.

Goldsmith, M., Sarov-Blat, L. & Livneh, Z., 2000. Plasmid-encoded *MucB* protein is a DNA polymerase (pol RI) specialized for lesion bypass in the presence of *MucA', RecA*, and *SSB*. *Proceedings of the National Academy of Sciences*, 97(21), pp.11227–11231.

González, J.M., Brown, B.J. & Carlton, B.C., 1982. Transfer of *Bacillus thuringiensis* plasmids coding for delta-endotoxin among strains of *B. thuringiensis* and *B. cereus*. *Proceedings of the National Academy of Sciences of the United States of America*, 79(22), pp.6951–6955.

Hale, L. et al., 2010. An efficient stress-free strategy to displace stable bacterial plasmids. *BioTechniques*, 48, pp.223–228.

Johnson, T.J. et al., 2007. Plasmid replicon typing of commensal and pathogenic *Escherichia coli* isolates. *Applied and Environmental Microbiology*, 73(6), pp.1976–1983.

Kobayashi, I., 2004. Genetic addiction: A principle of gene symbiosis in a genome. In B. E. Funnell & G. J. Phillips, eds. *Plasmid Biology*. Washington D.C.: American Society for Microbiology (ASM), pp. 105–144.

Lawley, T., Wilkins, B.M. & Frost, L.S., 2004. Bacterial conjugation in Gram-negative bacteria. In B. E. Funnell & G. J. Phillips, eds. *Plasmid Biology*. Washington D.C.: American Society for Microbiology (ASM), pp. 203–226.

Lehnherr, H. & Yarmolinsky, M.B., 1995. Addiction protein *Phd* of plasmid prophage P1 is a substrate of the *ClpXP* serine protease of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America*, 92(8), pp.3274–3277.

Lehnherr, H. et al., 1993. Plasmid addiction genes of bacteriophage P1: *doc*, which causes cell death on curing of prophage, and *phd*, which prevents host death when prophage is retained. *Journal of molecular biology*, 233(3), pp.414–428.

Liebana, E. et al., 2006. Longitudinal farm study of extended-spectrum β-lactamase-mediated resistance. *Journal of clinical microbiology*, 44(5), pp.1630–1634.

Lin, H. et al., 2005. In vitro characterization of salmochelin and enterobactin trilactone hydrolases *IroD, IroE*, and *Fes*. *Journal of the American Chemical Society*, 127(31), pp.11075–11084.

Maor-Shoshani, A. et al., 2000. Highly mutagenic replication by DNA polymerase V (*UmuC*) provides a mechanistic basis for SOS untargeted mutagenesis. *Proceedings of the National Academy of Sciences*, 97(2), pp.565–570.

Markowitz, V.M. et al., 2014. IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Research*, 42(Database issue), pp.D560–D567.

Ni, L. et al., 2009. The *Staphylococcus aureus* pSK41 plasmid-encoded *ArtA* protein is a master regulator of plasmid transmission genes and contains a RHH motif used in alternate DNA-binding modes. *Nucleic Acids Research*, 37(20), pp.6970–6983.

Overbeek, R. et al., 2014. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Research*, 42(Database issue), pp.D206–D214.

Pullinger, G.D. & Lax, A.J., 1992. A *Salmonella* dublin virulence plasmid locus that affects bacterial growth under nutrient-limited conditions. *Molecular Microbiology*, 6(12), pp.1631–1643.

Reuven, N.B. et al., 1999. The mutagenesis protein *UmuC* is a DNA polymerase activated by *UmuD′, RecA*, and *SSB* and is specialized for translesion replication. *Journal of Biological Chemistry*, 274(45), pp.31763–31766.

Reuven, P. & Eldar, A., 2011. Macromotives and microbehaviors: the social dimension of bacterial phenotypic variability. *Current Opinion in Genetics & Development*, 21(6), pp.759–767.

Roos, U., Harkness, R.E. & Braun, V., 1989. Assembly of colicin genes from a few DNA fragments. Nucleotide sequence of colicin D. *Molecular Microbiology*, 3(7), pp.891–902.

Rozen, S. & Skaletsky, H., 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods in molecular biology (Clifton, N.J.)*, 132, pp.365–386.

Rutherford, K. et al., 2000. Artemis: sequence visualization and annotation. *Bioinformatics (Oxford, England)*, 16(10), pp.944–945.

Sabri, M., 2006. A SitABCD homologue from an avian pathogenic *Escherichia coli* strain mediates transport of iron and manganese and resistance to hydrogen peroxide. *Microbiology*, 152(3), pp.745–758.

Salinger, N. et al., 2008. The plasmid-encoded regulator activates factors conferring lysozyme resistance on enteropathogenic *Escherichia coli* strains. *Applied and Environmental Microbiology*, 75(1), pp.275–280.

Sand, O., 2003. Phenotypic characterization of overexpression or deletion of the *Escherichia coli crcA, cspE* and *crcB* genes. *Microbiology*, 149(8), pp.2107–2117.

Sayeed, S. et al., 2000. The stability region of the large virulence plasmid of *Shigella flexneri* encodes an efficient postsegregational killing system. *Journal of Bacteriology*, 182(9), pp.2416–2421.

Stokes, M.O. et al., 2012. Detection and characterization of pCT-like plasmid vectors for blaCTX-M-14 in *Escherichia coli* isolates from humans, turkeys and cattle in England and Wales. *The Journal of antimicrobial chemotherapy*, 67(7), pp.1639–1644.

Tabuchi, A. et al., 1992. Autoregulation of the stability operon of IncFII plasmid NR1. *Journal of Bacteriology*, 174(23), pp.7629–7634.

Tabuchi, A. et al., 1988. Genetic organization and nucleotide sequence of the stability locus of IncFII plasmid NR1. *Journal of molecular biology*, 202(3), pp.511–525.

van Heel, A.J. et al., 2013. BAGEL3: automated identification of genes encoding bacteriocins and (non-)bactericidal posttranslationally modified peptides. *Nucleic Acids Research*, 41(Web Server issue), pp.W448–W453.

Villa, L. et al., 2010. Replicon sequence typing of IncF plasmids carrying virulence and resistance determinants. *Journal of Antimicrobial Chemotherapy*, 65(12), pp.2518–2529.

Willetts, N. & Skurray, R., 1980. The conjugation system of F-like plasmids. *Annual review of genetics*, 14, pp.41–76.

Chapter 5: Plasmid field study

5.1 Introduction

5.1.1 Plasmid communities

Disease-causing bacteria are a research priority, so it is understandable that the majority of plasmid population studies have focussed on clinically relevant plasmids, in particular those carrying virulence genes and antibiotic resistance (Johnson & Nolan 2009; Carattoli 2011; Marcade et al. 2008; Pallecchi et al. 2007; Dröge et al. 2000; Hopkins et al. 2006). In addition, many studies have relied on selectable markers, such as antibiotic or heavy metal tolerance, or a specific metabolic pathway, to isolate bacteria and plasmids. For example, Burton et al. (1982) isolated plasmids from polluted and unpolluted river sites, but selected for host bacteria able to tolerate pollution-mimicking media, and so excluded any strains (and therefore plasmids), which were not pollution tolerant. Although this probably produced a representative sample from the polluted site, it is less good for the unpolluted site. Few studies have investigated plasmid populations in a 'natural' setting, that is, a non-clinical environment without selection for specific traits.

One such study, which non-selectively examined plasmid populations from soil and fresh water found that overall plasmid presence was not evenly distributed across populations: in each environment, one host subgroup had very high plasmid prevalence (up to 100%), and contained multiple plasmids, whereas other strains carried no plasmids at all (Beilstein et al. 2008). This may indicate that some strains are more successful plasmid recipients and/or are better at retaining plasmids, corroborating the findings of Gordon (1992), who demonstrated that the number of native plasmids in a strain affects its ability to receive further plasmids: the presence of small plasmids increased the transfer rate, whereas presence of large plasmids, or indeed no plasmids, decreased it.

Sobecky and coworkers (Cook et al. 2001; Sobecky 2002; Sobecky et al. 1997) have studied marine plasmid communities extensively, finding variation in plasmid prevalence across sites and over time, and a bi-modal distribution of plasmid size, with small and large plasmids the most prevalent, echoing the results of Beilstein et al. (2008) and Gordon (1992).

pMLST (plasmid multi locus sequence typing) has been used to identify the prevalence of plasmids in livestock, and has demonstrated the presence of identical ESBL genes and plasmids in livestock and humans, suggesting the transfer of bacteria and plasmids through the food chain (Dierikx et al. 2012; Dierikx et al. 2010; Hordijk et al. 2013; Hall et al. 2011). However, these studies selected bacteria on the basis of resistance, thus excluding other host strains and plasmids. In addition, some sample sizes were small, and herd and host animal replication was low, giving little indication of the structure of the bacterial population (Hordijk et al. 2013; Dierikx et al. 2010).

5.1.2 Plasmids within hosts

Plasmids do not exist as isolated entities. Smalla & Sobecky (2002) state that "attempts to assess the HGP [horizontal gene pool] should ideally be combined with the study of structural and functional diversity of microbial communities", and Thomas (2004) concludes that future research must enable us to "understand better how plasmids interact within microbial communities". Plasmid populations can be seen as analogous to parasite populations in higher organisms, where it is recognised that host population structure can have a major impact on the parasites, and vice versa (Schmid-Hempel 2001). How plasmid population structure relates to the host bacterial population structure could provide insights into plasmid maintenance and host-plasmid coevolution, for example the relative importance of vertical vs. horizontal plasmid transmission (Harrison & Brockhurst 2012). Mechanisms of plasmid identification which dissociate plasmids from hosts (Section 5.1.3) clearly inhibit this type of community analysis, and to date few studies have examined natural plasmid and host communities.

5.1.3 Plasmid isolation strategies

Three principal methods are used to isolate and identify plasmids: i) endogenous, culture-based techniques; ii) exogenous mating-based techniques; and iii) molecular techniques (Smalla & Sobecky 2002). Endogenous plasmid isolation requires host cell culture: strains must be isolated and cultured to obtain clonal colonies from which plasmids can be extracted. This method has obvious limitations: bacteria which are difficult (or impossible) to culture in the laboratory will be ignored; the most prevalent strains in the environment will dominate, often overlooking rare strains and species; and in addition, selective media containing antibiotics are often required to isolate the desired bacteria (Smalla et al. 2005). However, the great advantage of endogenous isolation is that the host strain is known.

Exogenous plasmid isolation does not require bacterial culture. Also known as 'plasmid capture', this technique involves a marked 'capture' strain which acquires plasmids of interest via conjugation (Bale et al. 1988). Two variations exist: bi-parental mating, where a self-transmissible target plasmid transfers directly to the capture strain (Dahlberg et al. 1997); and tri-parental mating, where a second marked plasmid is mobilised by the target plasmid (Top et al. 1994). Bi-parental mating requires the target plasmid to carry a selectable marker, whereas tri-parental mating utilizes a secondary marked plasmid. However, the major limitation of these techniques is that they require the target plasmid to transfer to the capture strain, ruling out collection of non-transmissible or narrow host range plasmids. Moreover, these methods give no information about the original plasmid host.

Plasmids can also be identified using molecular techniques, for example: PCR based strategies such as replicon typing (Carattoli et al. 2005); hybridization probes (Couturier et al. 1988); pMLST (Garcia-Fernandez et al. 2008); and recently, plasmid motility-based identification (Garcillán-Barcia et al. 2011). These techniques have the advantage of being high throughput, and, like exogenous methods, do not require culture. They can even use environmental DNA extracts (Carattoli et al. 2005; Elsas & Bailey 2002; Sobecky et al. 1997).

Although this facilitates the collection of massive plasmid presence data sets, environmental DNA extracts have the major disadvantage of not retaining any individual plasmids for further study. Moreover, as with exogenous methods, these molecular approaches do not give information about the host bacterial population. Combination strategies, where molecular techniques are applied to endogenously or exogenously obtained plasmids, provide a balance between the high-throughput molecular techniques, and the specificity of plasmid extraction. Couturier probes (Couturier et al. 1988) have been successfully used for molecular analysis of endogenously isolated plasmids, allowing plasmids to be assigned a replicon (incompatibility) type (Sobecky et al. 1997).

5.1.4 Aims

The original aim of this chapter was to identify antibiotic resistance plasmids and corresponding susceptible 'cheat' plasmids in a natural bacterial population: enteric *E. coli* from cattle (see Chapter 4, Section 4.1.1 for an explanation of 'cheat' plasmids). Identification of 'cheat' plasmids required non-selective isolation methods, making this study an ideal opportunity to examine a natural plasmid population in a non-selective way: the only selection criterion for plasmid isolation was that the host must be *E. coli*.

This study combines endogenous, culture-based plasmid isolation with PCR based plasmid replicon identification and a novel method of host genotyping, to provide a picture of both the host and plasmid populations in a high-throughput manner. This chapter presents an overview of a very large data set: approximately 25000 *E. coli* were isolated, and to date 527 have been genotyped and screened for plasmids and antibiotic susceptibility. Such a substantial data set presents a unique opportunity to examine the ecology of a 'real' plasmid population. Fundamental questions about the coexistence of resistant and susceptible plasmids, plasmid and host diversity-abundance relationships, plasmid movement within populations, and population structuring of both host and plasmid are addressed.

5.2 Methods

5.2.1 Sample collection & isolation of E. coli

Bovine faecal samples (cowpats) were collected from 9 sites across Surrey and Berkshire: 4 farms (sites A, D, R & W) and 5 nature reserves with grazing cattle (sites CC, OC, RH, TM and WC). Sampling was conducted in a hierarchical manner: 10 cowpats were sampled from each site and 10 *E. coli* were screened from each pat (unless numbers were too low, in which case all available pats/ *E. coli* were screened). *E. coli* survival in cowpats in the field is good for up to 30 days, but fresh pats were selected when possible (van Kessel et al. 2007).

Approximately 2 g faecal matter was collected from each pat using sterile collecting tubes with spoon lids (Nunc™, Thermo Scientific, UK). Samples were returned to the laboratory where they were stored at 10°C until processing. All samples were processed with 24 hours of collection by homogenisation in 5 ml sterile 0.85% w/v NaCl followed by centrifugation at 500 g for 2 minutes to pellet the solid material. The supernatants were diluted 100-fold with sterile 0.85% w/v NaCl, and 100 μl was plated onto HiCrome™ Coliform Agar containing 5 mg l^{-1} novobiocin (Sigma, UK). Total coliform counts were recorded after overnight incubation at 37°C. Ten colonies per cowpat sample were re-streaked onto HiCrome™ plates to obtain clonal colonies. Overnight cultures in LB broth were inoculated from these clonal colonies for testing with Kovac's Reagent (Sigma). 10 μl Kovac's Reagent was added to 150 μl overnight culture; a cherry red colour confirms presence of *E. coli*. Isolated colonies were stored long term in 80% glycerol at -80°C.

5.2.2 DNA extraction & plasmid replicon typing

Overnight cultures were set up in 2 ml LB broth in 24 well plates. These were centrifuged at 5500 *g* for 5 minutes and the pellet resuspended in 0.5 ml molecular grade H_2O. This suspension was boiled for 10 minutes at 100°C and then centrifuged at 4500 *g* for 5 minutes (Johnson & Brown 1996). Supernatants were used as templates for replicon typing PCR and flagellin sequencing. Best

results were obtained when supernatants were used immediately, but they were also stored at -20°C.

Replicon typing PCR was used to identify plasmids in isolated *E. coli*. This method comprises 3 multiplex-PCR panels with 18 primer pairs recognising the major plasmid incompatibility groups of the *Enterobacteriaceae* (Carattoli et al. 2005; Johnson et al. 2007). PCR reagent concentrations were as follows: 0.4mM dNTPs, 0.625 units Taq (Qiagen) and 2.4 µM of each primer in the multiplex panel. 0.5 µl boiled lysate template was added to each 24.5 µl reaction mix. Cycling conditions were: initial denaturation, 95°C, 5 minutes, followed by 30 cycles of denaturation (95°C, 30 seconds), annealing (60°C, 30 seconds) and extension (72°C, 90 seconds) and a final extension, step of 72°C for 5 minutes. PCR product sizes indicate the presence of specific plasmid replicon(s).

5.2.3 E. coli genotyping

E. coli isolates were genotyped by sequencing the variable region of the flagellin gene, *fliC*. This method is an adaptation of previous genotyping methods using whole gene sequencing (Wang et al. 2006), and PCR-RFLP (restriction fragment length polymorphism) (Machado et al. 2000). Whole gene sequencing was not practicable for such a large sample size (>500 isolates) given the length of the *fliC* gene (between 0.9-2.5Kb). A novel genotyping method involving sequencing the variable region of *fliC* was developed. The *fliC* gene was amplified using the PCR-RFLP primers described by Machado et al. (2000)(Table 5.1). Standard PCR reagents can be seen in Chapter 2, Section 2.2.1. Cycling conditions were as follows: initial denaturation, 95°C, 2 minutes, followed by 30 cycles of denaturation (95°C, 40 seconds), annealing (50°C, 30 seconds) and extension (72°C, 120 seconds) and a final extension step of 72°C for 3 minutes. PCR products were cleaned up using Exo1 (Fisher Scientific) and TSAP (Promega UK, Southamption, UK) both at final concentration 0.045 units μl^{-1}. Clean up mixes were incubated at 37°C for 30 minutes, followed by 80°C for 15 minutes. Sequencing reactions were conducted using BigDye® terminator v3.1 cycle sequencing kit (Life Technologies Ltd. Paisley, UK) using the *fliC 2* primer only.

Sequencing cycling conditions were as follows: 96°C for 1 minute, followed by 25 cycles of: 96°C for 10 seconds, 50°C for 5 seconds, 60°C or 4 minutes. Sequencing reactions were cleaned up using ethanol/sodium acetate precipitation. Sequencing runs were conducted by the Department of Zoology, University of Oxford, UK. Trimmed sequences were searched against the whole flagellin gene sequences of Wang et al. (2003) using Geneious (Biomatters 2012). Sequences were assigned a flagellin type (H-antigen type) when they had at least 200 bp of high identity (>95%) with one *fliC* sequence. Where several matches were made the match with highest identity was selected. This method allowed 97% sequences to be assigned a flagellin type.

Primer	**Sequence (5'-3')**	**Size (nt)**
fliC 1	CAAGTCATTAATACMAACAGCC	22
fliC 2	GACATRTTRGAVACTTCSGT	20

Table 5.1: Primers for flagellin typing. Primers were taken from Machado et al. (2000), and include 'wobble' bases: M (A/C), R (A/G), V (G/A/C), S (G/C).

E. coli H-antigen types were placed into one of four subgroups as defined by Wang et al. (2003). These are based on H-antigen groupings of variable region of the flagellin gene (V), which was sequenced (Table 5.2). Wang et al. (2003) failed to categorise the H16 group based on its V region sequence, but place it in the Ec1b group based on the conserved C1 and C2 regions. They suggest that this failure is due to recent recombination in the V region of the gene.

Ec1a			**Ec1b**		**Ec2**	**Se2**
H6	H18	H32	H4	H38	H8	H2
H7	H19	H37	H5	H39	H11	H48
H12	H20	H41	H10	H42	H21	
H14	H28	H49	H16	H56	H27	
H15	H31		H33			

Table 5.2: *E. coli* H-antigen subgroupings

5.2.4 Antibiotic susceptibility screening

The susceptibility of *E. coli* isolates to ampicillin, cefotaxime and chloramphenicol was assessed by disc diffusion assays. MIC breakpoints and interpretation of zone diameters were taken from The British Society for Antimicrobial Chemotherapy Guidelines version 11.1 (May 2012). Susceptibility testing methods were adapted from Andrews et al. (2005). Overnight cultures in 1.5 ml LB broth in 24 well plates were initiated from single colonies. Plates were inoculated with bacteria by spreading in three directions with a sterile swab. Six antibiotic discs were placed at uniform intervals using sterile forceps (Table 5.3). The zones of inhibition were measured after overnight incubation at 37°C, using a disc diameter template. Two doses were used for each antibiotic, and strains were determined to be resistant if they scored positively with both discs.

Antibiotic	**Doses (μg ml $^{-1}$)**		**Clearance zone diameter (mm)**
Ampicillin	2	10	15
Cefotaxime	5	30	30
Chloramphenicol	10	30	21

Table 5.3: Antibiotic disc doses and diameters of clearance.

5.2.5 Statistical analysis

All statistical analyses were conducted in *R* (R Core Team 2013). Plasmid and H-antigen diversity were calculated using the Shannon diversity index (Spellerberg & Fedor 2003) using the *Vegan* package in *R* (Oksanen et al. 2013). Linear regressions were carried out to assess the affect of E. coli abundance on both plasmid and H-antigen diversity, and also to compare both diversity measures. In addition, a mixed model incorporating 'site' as a random effect was conducted using the *R* package *nlme* (Pinheiro et al. 2007). Restricted maximum likelihood (REML) lme models were converted to maximum likelihood (ML) models to allow model comparison using ANOVA. Plasmid replicon distribution across subgroups was assessed using a mosaic plot, produced using the *vcd* package

(Hornik et al. 2006). Hierarchical *F*-statistics were calculated using *HierFstat* (Goudet 2004; de Meeûs & Goudet 2007).

5.3 Results

5.3.1 Antibiotic resistance is rare

In total 24483 colonies were isolated from 84 cowpats across 9 sites. Of these, 527 isolates were genotyped and screened for plasmid replicons and antibiotic susceptibility. Somewhat surprisingly, very low levels of antibiotic resistance were found: of the 527 isolates just 14 (less that 3%) were resistant to any of the antibiotics tested: 9 isolates were ampicillin resistant, 4 were chloramphenicol resistant, and just 1 was resistant to cefotaxime. No multi-resistant isolates were found. Resistant isolates were found in 5 of the 9 sites, with sites A, CC, OC and W resistance-free. The absence of resistance from sites A and CC may be an artefact of low *E. coli* recovery from these sites (see Section 5.3.2 below), but despite this the number of resistant isolates is much lower than was anticipated. The rarity of antibiotic resistance made it impossible to test for the co-occurrence of resistant and susceptible bacteria, and therefore the presence of 'cheat' plasmids.

5.3.2 E. coli abundance is a poor predictor of plasmid diversity

It seems logical that a larger *E. coli* population may support a greater variety of plasmid replicons. In order to test this, regression was conducted with total *E. coli* abundance against plasmid replicon diversity (Shannon diversity). When the data was pooled by site, *E. coli* abundance was a moderately good predictor of plasmid replicon diversity ($F_{(1, 7)} = 5.61$, $R^2 = 0.3656$, $p = 0.0497$; Figure 5.1 A, dashed line). However, this diversity-abundance relationship appears to be skewed by sites A and CC, where low numbers of *E. coli* were isolated (Figure 5.1 A, open circles). When these sites were excluded, the diversity-abundance relationship was not significant ($F_{(1, 5)} = 3.643$, $R^2 = 0.3058$, $p = 0.1146$; Figure 5.1 A, solid line). When data was pooled by individual cowpat, *E. coli* abundance explains a smaller amount of the variance in plasmid replicon diversity, but this relationship is highly significant ($F_{(1, 82)} = 30.45$, $R^2 = 0.2619$, $p = 3.90 \times 10^{-7}$, Figure 5.1 B).

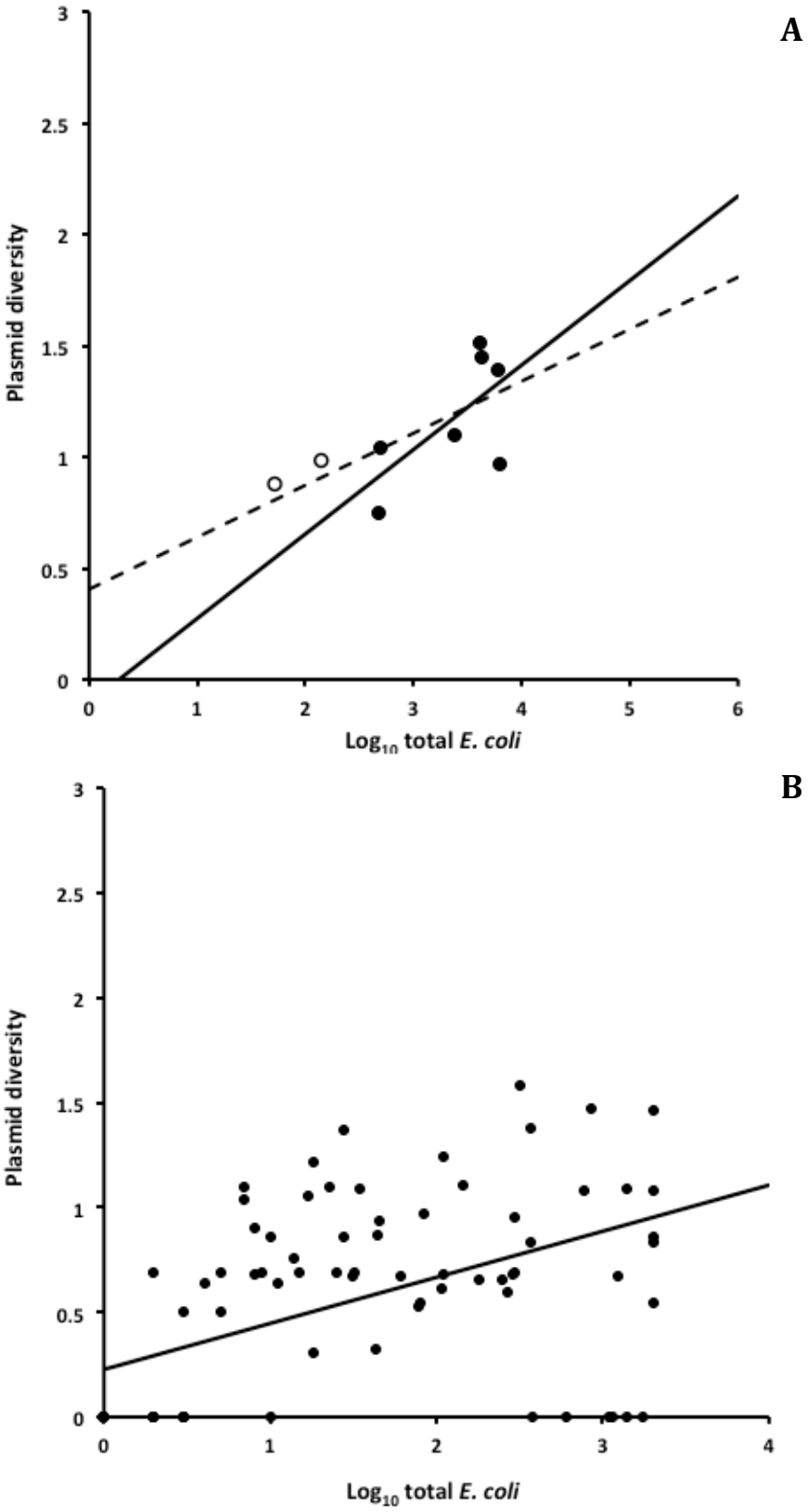

Figure 5.1: Plasmid replicon diversity vs. *E. coli* abundance. (A) Site level. When all sites are included, *E. coli* abundance is a good predictor of plasmid diversity ($F_{(1, 7)}$ = 5.61, R^2 = 0.3656, p= 0.0497, dashed line). However, when sites with low *E. coli* abundance were excluded (open circles), this relationship is not significant ($F_{(1, 5)}$ = 3.643, R^2 = 0.3058, p = 0.1146, solid line). **(B)** Pat level. *E. coli* abundance is a reasonably good predictor of plasmid replicon diversity at the pat level ($F_{(1, 82)}$ = 30.45, R^2 = 0.2619, p = 3.90 x 10^{-7}). Diversity is Shannon diversity.

The weak correlation between diversity and abundance could be explained by differences between sites. Pat level plasmid replicon diversity and total *E. coli* were modelled using a mixed effects model with site as the random effect (AIC 91.32). Figure 5.2 shows variation in the slope of this relationship between sites: there is generally a positive correlation between plasmid replicon diversity and *E. coli* abundance, with the exception of sites OC and R, which show slight negative trends. Despite these differences, there is no significant decline in the model's explanatory power when the random effect is excluded (AIC 90.87, p = 0.213), indicating that the random effect of site is not important. Overall, these data suggest that *E. coli* abundance does not significantly affect plasmid replicon diversity at the site level, but is significant at the pat level.

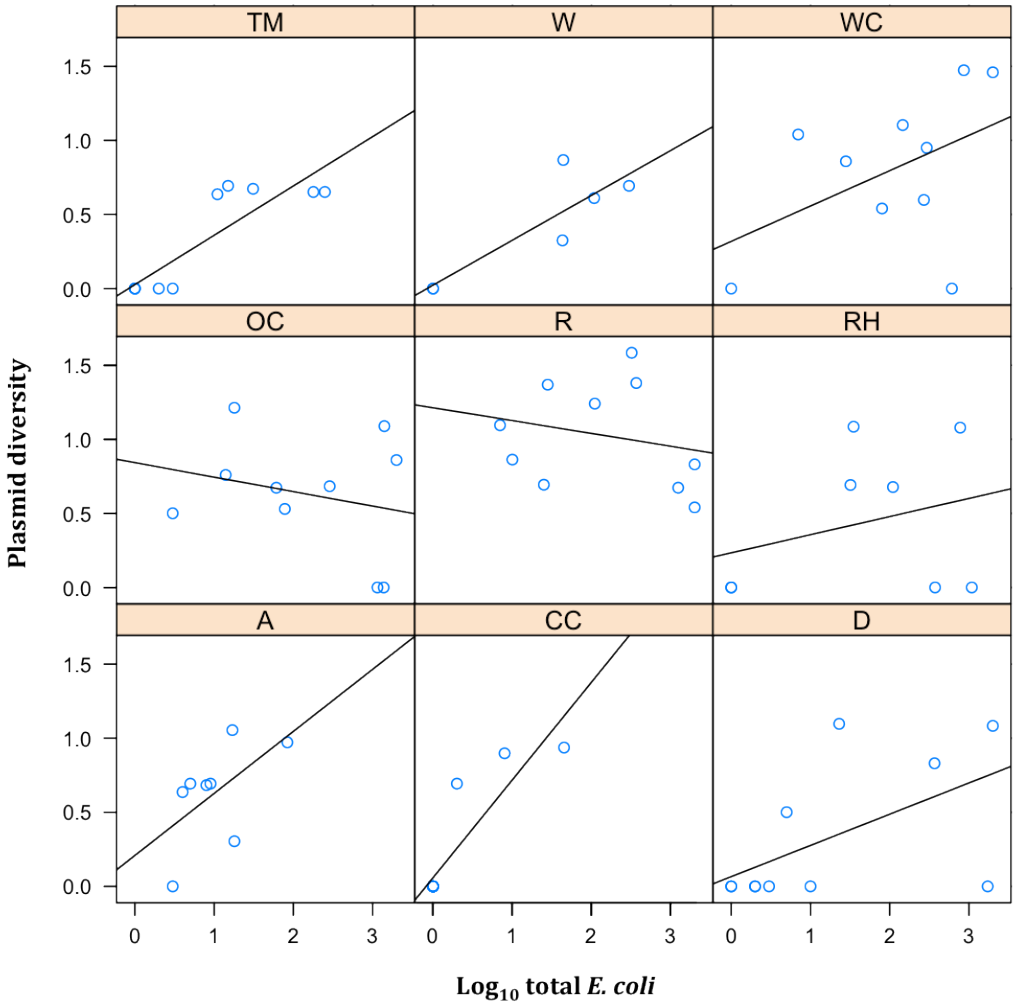

Figure 5.2: Pat level plasmid replicon diversity vs *E. coli* abundance for each of the nine sites. Each panel represents the pat level data for one site. Total *E. coli* was transformed by the addition of the minimum detectable value to allow zeros to be plotted.

5.3.3 *Replicons are not evenly distributed across E. coli subgroups*

The majority of *E. coli* isolates contained at least one plasmid replicon; just 26 % (136 isolates) had no detectable plasmids. Of the 18 possible replicons picked up by the screen 8 were detected. There was one clear dominant type, FIB, which was found in 352 isolates (67%). Of the other replicon types identified, FIA and P were common (> 100 isolates); FIC and I1 were moderately common (> 30 isolates); and B/O, K/B and Y were rare (< 10 isolates). The full plasmid and H-antigen data sets can be seen in Appendix 2.

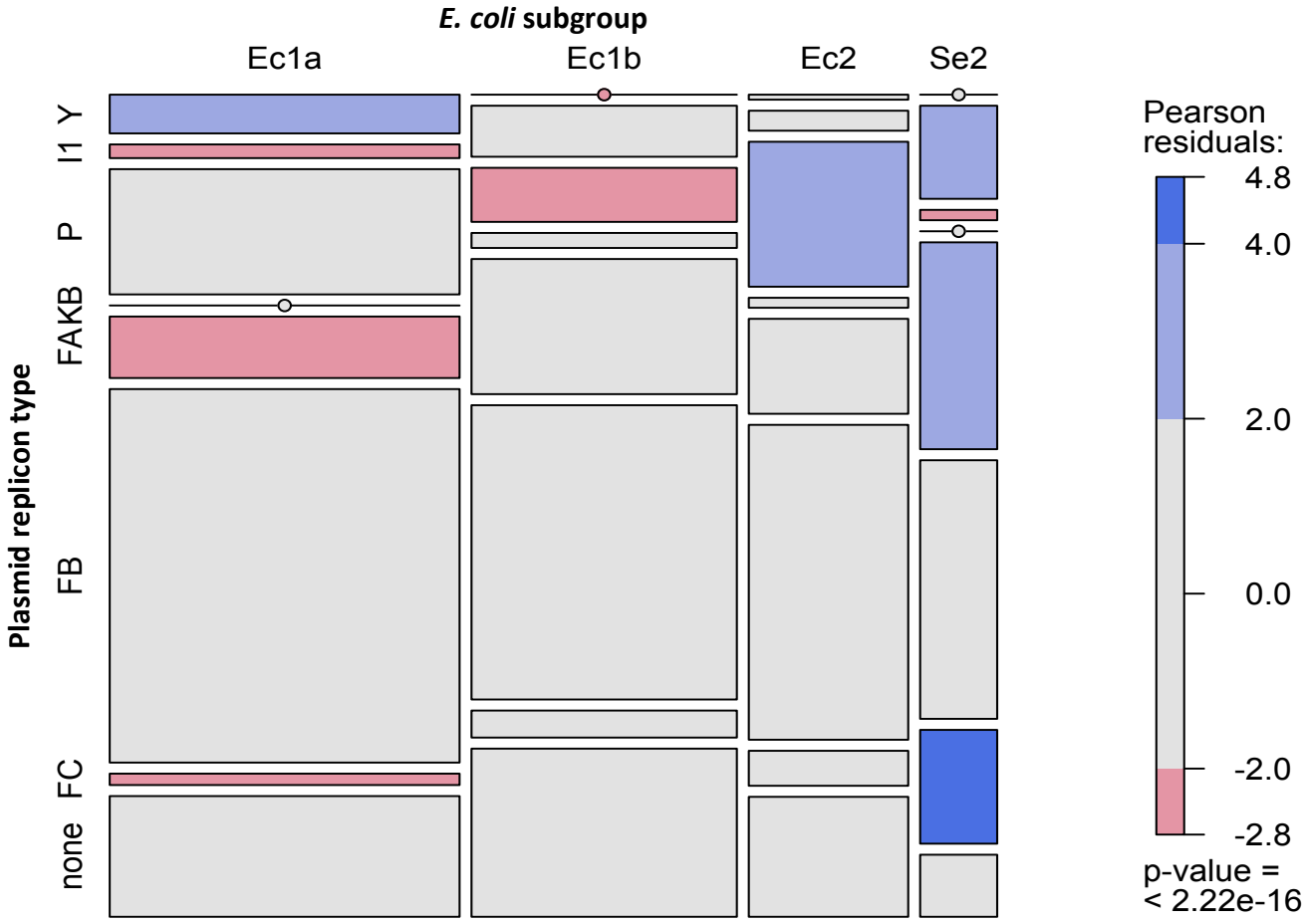

Figure 5.3: Mosaic plot showing plasmid replicons within each *E. coli* subgroup. Tiles represent the number or replicons of each type (y axis) present in each subgroup (x axis). Bullet points indicate zero. Replicons: FIA (FA), FIB (FB), FIC (FC), P, I1, Y, pooled K/B & B/O (KB) and none. Shading shows Pearson residuals, which indicate the departure of each tile from independence; a residual > 2 or < -2 implies a departure significant at the 95% level, and a residual > 4 or < -4 implies a departure significant at 99.99% level. Therefore, grey tiles indicate no significant variation from independence, red tiles indicate significantly fewer replicons of that type than independence predicts, and blue tiles indicate significantly more. Ec1a, n = 327; Ec1b, n = 248; Ec2, n = 149; Se2, n = 72 (see Appendix 2 for data set).

The mosaic plot (Figure 5.3) highlights non-even distribution of replicons across *E. coli* subgroups. Significantly more replicons than independence would expect were found for: Y in Ec1a; P in Ec2; I1, FIA and FIC in Se2. The replicon distribution for subgroup Se2 is particularly striking, especially as this is the smallest subgroup (n = 72). 3 replicon types are overrepresented in Se2, including FIC, which has the only Pearson residual of > 4, indicating significance at the 99.99% level. Significantly fewer replicons than independence predicts were found for: I1, FIA and FIC in Ec1a; Y and P in Ec1b; P in Se2. Curiously, the largest subgroup, Ec1a (n = 327), shows the most underrepresentation of plasmid replicons. Distribution of *E. coli* with no plasmids is even across all four subgroups, as are the distributions of the most and least common plasmid types, FIB (n = 352) and pooled K/B and B/O (KB, n = 7). The colourfulness of the mosaic plot clearly demonstrates that plasmids are not evenly distributed across *E. coli* subgroups in this study.

5.3.4 E. coli diversity can predict plasmid diversity

E. coli genotype diversity was calculated as Shannon diversity of H-antigen types at the site level. Figure 5.4 shows that this H-antigen diversity correlates significantly with plasmid replicon diversity. When all sites were included, H-antigen diversity is a reasonably good predictor of plasmid diversity ($F_{(1, 7)}$ = 6.518, R^2 = 0.4082, p = 0.0379, Figure 5.4, dashed line). When the low *E. coli* abundance sites, A and CC, were excluded (Figure 5.4, open circles) this relationship has greater explanatory power and significance ($F_{(1, 5)}$ = 9.498, R^2 = 0.5862, p = 0.0274, solid line). It is interesting to note that *E. coli* diversity, not abundance, affects the diversity of plasmids in the population.

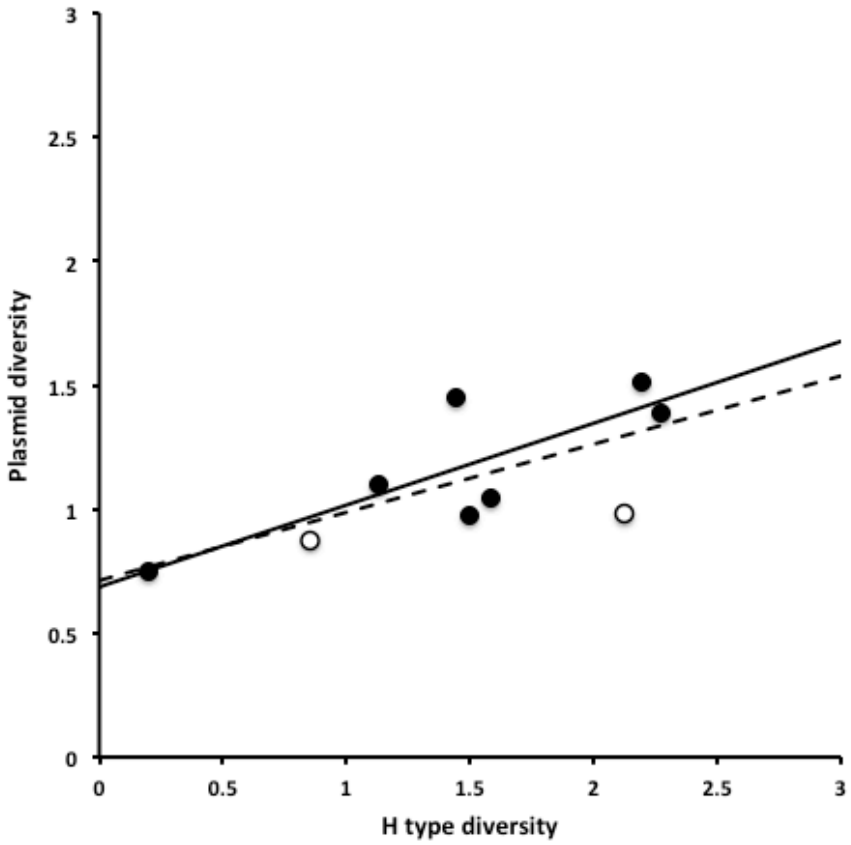

Figure 5.4: Plasmid replicon diversity correlates with H-antigen diversity. When all sites are included, H-antigen diversity is a reasonably good predictor of plasmid diversity ($F_{(1, 7)}$ = 6.518, R^2 = 0.4082, p = 0.0379, dashed line). When the low *E. coli* abundance sites are excluded (open circles) this relationship becomes more significant ($F_{(1, 5)}$ = 9.498, R^2 = 0.5862, p = 0.0274, solid line). Both are Shannon diversity at site level.

5.3.5 E. coli diversity & population structure

In addition to studying the plasmid population, it is also important to consider the host *E. coli* population. Regression was conducted with total *E. coli* abundance against H-antigen diversity at the site level. *E. coli* abundance was found to be a poor predictor of H-antigen diversity, both when low abundance sites are included ($F_{(1,7)}$ = 1.1002, R^2 = 0.0124, p= 0.3291, Figure 5.5 dashed line) and when they are excluded ($F_{(1,5)}$ = 3.5051, R^2 = 0.2945, p = 0.1201, Figure 5.5 solid line).

It was observed that within each pat most *E. coli* were of the same H-antigen type. Hierarchical *F*-statistics were calculated to assess the contribution of sites, and pats within sites to the *E. coli* population structure: how much genotype variation (different H-antigen types) can be explained by each level of the hierarchy (pat, site). Both pat within site ($F_{pat/site} = 0.457$, $p = 1 \times 10^{-4}$), and site ($F_{site/total} = 0.197$, $p = 1 \times 10^{-4}$) were found to have strong effects on the population structure of *E. coli* (10 000 permutations). In other words, 20% of the observed genotype variation can be explained by site, and 46% of this variation can be explained by pat within site, corresponding to great and very great genetic differentiation respectively (Holsinger & Weir 2009; Wright 1978).

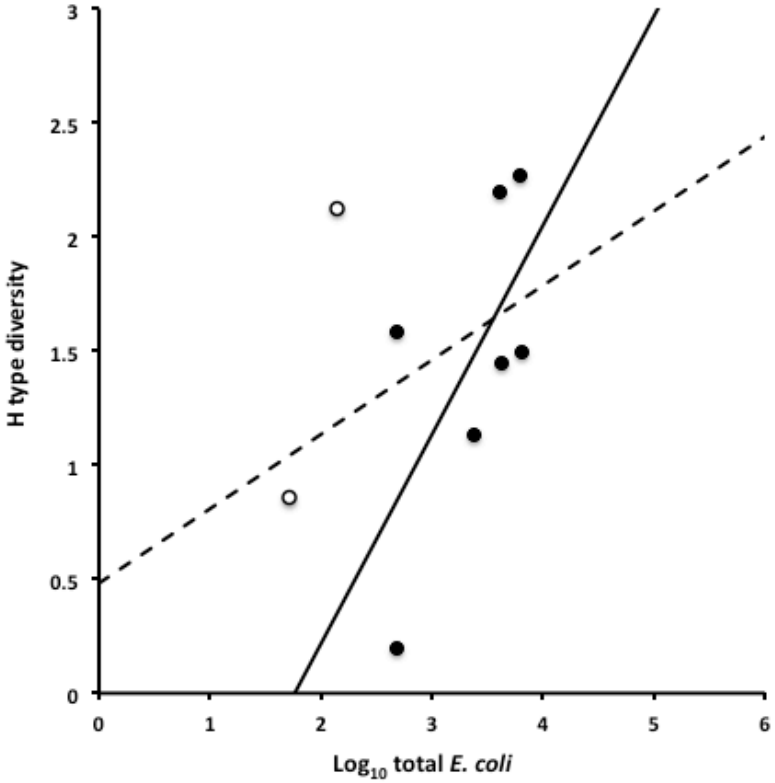

Figure 5.5: H-antigen diversity vs. *E. coli* abundance. *E. coli* abundance is a poor predictor of H-antigen diversity, both when low abundance sites are included ($F_{(1,7)} = 1.1002$, $R^2 = 0.0124$, $p = 0.3291$, dashed line) and when they are excluded ($F_{(1,5)} = 3.5051$, $R^2 = 0.2945$, $p = 0.1201$, solid line). H-antigen diversity is Shannon diversity per site.

5.4 Discussion

5.4.1 Antibiotic resistance and the food chain

It has been proposed that farm animals may be important carriers of, and reservoirs for, ESBL genes (Carattoli 2008). However, data on the prevalence of ESBL producing bacteria in food or food-animals is highly variable, ranging from 0.2% to 40% (Liebana et al. 2013). This difference is highlighted by the contrast between data in this study and a previous study of faecal samples from dairy cattle by Sawant et al. (2007): < 3 % of *E. coli* were resistant to ampicillin, chloramphenicol or cefotaxime in my study, whereas 48 % and 20 % of *E. coli* were resistant to ampicillin and chloramphenicol respectively in Sawant et al.'s work. It is of course possible that the *E. coli* sampled in this study carry resistance to antibiotics other than those tested. However, the lack of ESBL resistance is noteworthy as it may indicate that resistance in livestock is not as prevalent as is supposed. Five of the nine sites in the study were small grazing herds kept on nature reserves. It is highly likely that these cattle have had very low exposure to antibiotics, and the low levels of resistance found suggest that this means of cattle rearing may reduce the prevalence of resistance. If this is the case, it provides hope that the implementation of control measures could have a positive impact on the spread of resistance: all is not lost. This data suggests that not all *E. coli* populations are rife with antibiotic resistance. The spread of antibiotic resistant organisms through the food chain is considered to be a significant risk to public health, but if resistance in livestock can be managed effectively, this risk may be minimised (Liebana et al. 2013).

5.4.2 High numbers IncF plasmids

Replicon typing detected a large number of plasmids, with 74 % of *E. coli* isolates testing positive for at least one replicon. Perhaps unsurprisingly, the most commonly detected replicons were IncF (FIA, FIB and FIC). IncF plasmids are one of the most common plasmid types in the *Enterobacteriaceae*, are found across a range of host animals (Johnson et al. 2007), and have previously been identified as the most common plasmids in enteric *E. coli* from cattle (Hordijk et

al. 2013) Interestingly for this thesis, the plasmids sequenced in Chapter 4 were also IncF, and were isolated from *E. coli* from a cattle farm. IncF plasmids are often associated with virulence genes and antibiotic resistance, in particular ESBLs (Villa et al. 2010). None of the plasmids sequenced in Chapter 4 carry ESBL genes, and considering the lack of resistance observed in this study, it seems unlikely that these plasmids do either. Further sequencing of a selection of these plasmids would give a fuller picture of the non-resistant IncF plasmid population, complementing the work in Chapter 4, and providing insights into the differences between these plasmids and their resistant counterparts. It should be noted that previous studies have found IncF replicon probes to be the most successful in detecting plasmids from wild strains, with other probes sometimes having difficulty detecting plasmids (Hales et al. 1992; Kobayashi & Bailey 1994). Plasmid DNA extraction and sequencing may therefore identify additional plasmids missed due to this detection issue, as well as small plasmids which might not have detectable replicons.

5.4.3 Host diversity determines plasmid diversity

Perhaps the most interesting finding in this study is that host abundance has little effect on plasmid replicon diversity, whereas host diversity has a significant effect (although some pat-level effect was seen). This finding, that it is not the number of available hosts but their diversity that affects replicon diversity suggests the association of plasmid replicons with specific host types, and plasmid spread by vertical transmission rather than conjugation. The observation that plasmid replicons are not evenly distributed across *E. coli* subgroups - that certain replicons are over- or underrepresented in certain subgroups - also supports the idea that specific plasmid replicons are associated with specific hosts in this population. Moreover, the structure of the host *E. coli* population, where *E. coli* within one pat (animal) are mostly of one H-antigen type, would support plasmid persistence via vertical transmission, as opportunities for conjugation within the gut would be limited without high segregation rates (see Section 5.4.4).

The data suggesting strong associations of plasmid replicons with hosts could indicate coevolution of plasmids and hosts through altered conjugation rate, loss of genes or changes in gene expression (Slater et al. 2008; Harrison & Brockhurst 2012). It would be interesting to ascertain whether the plasmids in this population are conjugative, whether host strains can be cured of plasmids (and what effect this has), and what the rates of segregation are. This would give further information as to whether these plasmids are moving towards obligate symbiosis with their host, or whether the population is just so clonal that there is no chance of horizontal transfer. In addition, this data only gives information at the level of the replicon: it is not possible to say whether the plasmids within one pat or *E. coli* genotype are identical. Further plasmid sequencing, or pMLST, available for IncF and IncI plasmids (Villa et al. 2010; Garcia-Fernandez et al. 2008), would allow finer resolution of the plasmid population.

In addition, sequencing of plasmids would allow plasmid recombination to be examined. As this study identified only plasmid replicons, it is not possible to assess whether these replicons represent independent plasmids, whether these plasmids are the same, or whether several *rep* genes may be carried on the same plasmid due to recombination. Co-occurrence of several replicons in the same host could provide the opportunity for plasmids to recombine, generating new plasmid variants. In particular, IncF plasmids are known to be large and often carry multiple rep genes (as demonstrated in Chapter 4 of this thesis). Co-occurring FIA, FIB and FIC replicons may in fact represent large multi-replicon recombinant plasmids, rather than several small plasmids.

5.4.4 E. coli population structure

Host population structure is recognized as a key determinant of genetic diversity in parasites (van den Broeck et al. 2014). As plasmids can be considered to act as parasites, it follows that the population structure of host cells may also have a large influence of plasmid genetic diversity. It is therefore vital to consider the host population structure when examining plasmid populations. In this study, many pats appear to contain clonal *E. coli* populations: all isolates from a given

pat have the same H-antigen type. We can assume that each pat is independent (from one animal), and from this we can infer that the *E. coli* population within one cow is largely clonal. In fact, this may be an underestimate of between-animal variation, as multiple pats from one animal may have been sampled. A previous study (with a much smaller sample size, n = 99; Sawant et al. 2007) found that multiple animals at the same farm carried *E. coli* of the same subgroup, but that subgroups were not found on more than one farm. Unfortunately, despite using direct animal sampling, it is not possible to assess the clonality of *E. coli* populations within animals in this study, as only 1-2 *E. coli* isolates were tested per animal. However, other studies have demonstrated that the enteric *E. coli* populations of cattle consist of few dominant *E. coli* subtypes and many transient subtypes, which may appear only once (Anderson et al. 2006; Jenkins et al. 2003). Human enteric *E. coli* have also shown a similar pattern (Caugant et al. 1981). My data fits well with this model of one dominant and several less common *E. coli* subtypes, and the larger numbers of pats and *E. coli* sampled adds weight to the conclusion, although it is not possible to say whether the less common subtypes are transient as all the samples were collected at one time point.

A highly clonal, low diversity, *E. coli* population within an animal has implications for the plasmid population. As discussed in Section 5.4.3 above, plasmid diversity correlates strongly with host diversity, so that a clonal host population is likely to have a clonal plasmid population. However, transient or rare *E. coli* subtypes could introduce new plasmids to the system via conjugation. Indeed Caugant et al. (1981) found changes in the plasmid profiles of the two dominant *E. coli* subtypes in a human gut over the course of their study.

5.4.5 Conclusions and outlook

This study presents some fundamental ecology of enteric *E. coli* and plasmids from cattle. It has found surprisingly low levels of resistance, giving hope that alternative means of cattle rearing may lead to lower prevalence of antibiotic resistance. It has also demonstrated the link between host *E. coli* diversity and

plasmid diversity, and found that *E. coli* populations are highly structured. Further sequencing of plasmids would allow us to build a fuller picture of the plasmid population, in particular IncF plasmids, complementing the work in Chapter 4 of this thesis. DNA extraction and sequencing may also uncover small plasmids that were missed by replicon typing, and enable more accurate analysis of plasmid mobility.

References:

Anderson, M.A., Whitlock, J.E. & Harwood, V.J., 2006. Diversity and distribution of *Escherichia coli* genotypes and antibiotic resistance phenotypes in feces of humans, cattle, and horses. *Applied and Environmental Microbiology*, 72(11), pp.6914–6922.

Andrews, J.M., 2005. BSAC standardized disc susceptibility testing method (version 4). *The Journal of antimicrobial chemotherapy*, 56(1), pp.60–76.

Bale, M.J., Day, M.J. & Fry, J.C., 1988. Novel method for studying plasmid transfer in undisturbed river epilithon. *Applied and Environmental Microbiology*, 54(11), pp.2756–2758.

Beilstein, F., Battermann, A. & Dreiseikelmann, B., 2008. Plasmid incidence in the subgroups of two bacterial communities. *Microbiological Research*, 163(3), pp.350–353.

Biomatters, 2012. Geneious version 5.2. Available at: http://www.geneious.com/.

Burton, N.F., Day, M.J. & Bull, A.T., 1982. Distribution of bacterial plasmids in clean and polluted sites in a South Wales river. *Applied and Environmental Microbiology*, 44(5), pp.1026–1029.

Carattoli, A., 2008. Animal reservoirs for extended spectrum β-lactamase producers. *Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, 14(s1), pp.117–123.

Carattoli, A., 2011. Plasmids in Gram negatives: Molecular typing of resistance plasmids. *International Journal of Medical Microbiology*, 301(8), pp.654–658.

Carattoli, A. et al., 2005. Identification of plasmids by PCR-based replicon typing. *Journal of microbiological methods*, 63(3), pp.219–228.

Caugant, D.A., Levin, B.R. & Selander, R.K., 1981. Genetic diversity and temporal variation in the *E. coli* population of a human host. *Genetics*, 98(3), pp.467–490.

Cook, M.A. et al., 2001. Endogenous isolation of replicon probes for assessing plasmid ecology of marine sediment microbial communities. *Microbiology*, 147(8), pp.2089–2101.

Couturier, M. et al., 1988. Identification and classification of bacterial plasmids. *Microbiological reviews*, 52(3), p.375-395.

Dahlberg, C. et al., 1997. Conjugative plasmids isolated from bacteria in marine environments show various degrees of homology to each other and are not closely related to well-characterized plasmids. *Applied and Environmental Microbiology*, 63(12), pp.4692–4697.

de Meeûs, T. & Goudet, J., 2007. A step-by-step tutorial to use HierFstat to analyse populations hierarchically structured at multiple levels. *Infection, Genetics and Evolution*, 7(6), pp.731–735.

Dierikx, C. et al., 2012. Extended-spectrum-β-lactamase- and AmpC- β -lactamase-producing *Escherichia coli* in Dutch broilers and broiler farmers. *The Journal of antimicrobial chemotherapy*, 68(1), pp.60–67.

Dierikx, C. et al., 2010. Increased detection of extended spectrum β-lactamase producing *Salmonella enterica* and *Escherichia coli* isolates from poultry. *Veterinary Microbiology*, 145(3-4), pp.273–278.

Dröge, M., Pühler, A. & Selbitschka, W., 2000. Phenotypic and molecular characterization of conjugative antibiotic resistance plasmids isolated from bacterial communities of activated sludge. *Molecular and General Genetics*, 263(3), pp.471–482.

Elsas, J.D. & Bailey, M.J., 2002. The ecology of transfer of mobile genetic elements. *FEMS Microbiology Ecology*, 42(2), pp.187–197.

Garcia-Fernandez, A. et al., 2008. Multilocus sequence typing of IncI1 plasmids carrying extended-spectrum β -lactamases in *Escherichia coli* and *Salmonella* of human and animal origin. *Journal of Antimicrobial Chemotherapy*, 61(6), pp.1229–1233.

Garcillán-Barcia, M.P., Alvarado, A. & La Cruz, De, F., 2011. Identification of bacterial plasmids based on mobility and plasmid population biology. *FEMS Microbiology Reviews*, 35(5), pp.936–956.

Gordon, D.M., 1992. Rate of plasmid transfer among *Escherichia coli* strains isolated from natural populations. *Journal of General Microbiology*, 138(1), pp.17–21.

Goudet, J., 2004. hierFstat, a package for R to compute and test hierarchical *F*-statistics. *Molecular Ecology Resources*, 5(1), pp.184–186.

Hales, B.A. et al., 1992. The large plasmids found in enterohemorrhagic and enteropathogenic *Escherichia coli* constitute a related series of transfer-defective Inc F-IIA replicons. *Plasmid*, 28(3), pp.183–193.

Harrison, E. & Brockhurst, M.A., 2012. Plasmid-mediated horizontal gene transfer is a coevolutionary process. *Trends in Microbiology*, 20(6), pp.262–267.

Holsinger, K.E. & Weir, B.S., 2009. Genetics in geographically structured populations: defining, estimating and interpreting FST. *Nature Reviews: Genetics*, 10(9), pp.639–650.

Hopkins, K.L. et al., 2006. Replicon typing of plasmids carrying CTX-M or CMY beta-lactamases circulating among *Salmonella* and *Escherichia coli* isolates. *Antimicrobial Agents and Chemotherapy*, 50(9), pp.3203–3206.

Hordijk, J. et al., 2013. Cross-sectional study on prevalence and molecular characteristics of plasmid mediated ESBL/AmpC-producing *Escherichia coli* isolated from veal calves at slaughter *PloS one*, 8(5), pp.e65681.

Hornik, K., Zeileis, A. & Meyer, D., 2006. The strucplot framework: Visualizing multi-way contingency tables with vcd. *Journal of Statistical Software*, 17(3), pp.1–48.

Jenkins, M.B. et al., 2003. Putative temporal variability of *Escherichia coli* ribotypes from yearling steers. *Journal of environmental quality*, 32(1), pp.305–309.

Johnson, J.R. & Brown, J.J., 1996. A novel multiply primed polymerase chain reaction assay for identification of variant papG genes encoding the Gal (α1–4) Gal-binding papG adhesins of *Escherichia coli. Journal of Infectious Diseases*, 173(4), pp.920–926.

Johnson, T.J. & Nolan, L.K., 2009. Pathogenomics of the virulence plasmids of *Escherichia coli*. *Microbiology and molecular biology reviews : MMBR*, 73(4), pp.750–774.

Johnson, T.J. et al., 2007. Plasmid replicon typing of commensal and pathogenic *Escherichia coli* isolates. *Applied and Environmental Microbiology*, 73(6), pp.1976–1983.

Kobayashi, N. & Bailey, M.J., 1994. Plasmids isolated from the sugar beet phyllosphere show little or no homology to molecular probes currently available for plasmid typing. *Microbiology (Reading, England)*, 140 (Pt 2), pp.289–296.

Leverstein van-Hall, M.A. et al., 2011. Dutch patients, retail chicken meat and poultry share the same ESBL genes, plasmids and strains. *Clinical Microbiology and Infection*, 17(6), pp.873–880.

Liebana, E. et al., 2013. The public health risks of enterobacterial isolates producing extended-spectrum β-lactamases (ESBL) or AmpC β-lactamases in food and food-producing animals: An EU perspective of epidemiology, analytical methods, risk factors and control options. *Clinical Infectious Diseases*, 56(7), pp.1030–1037.

Machado, J., Grimont, F. & Grimont, P.A., 2000. Identification of *Escherichia coli* flagellar types by restriction of the amplified fliC gene. *Research in microbiology*, 151(7), pp.535–546.

Marcade, G. et al., 2008. Replicon typing of plasmids in *Escherichia coli* producing extended-spectrum β-lactamases. *The Journal of antimicrobial chemotherapy*, 63(1), pp.67–71.

Oksanen, J. et al., 2013. Vegan: community ecology package. Available at: http://CRAN.R-project.org/package=vegan.

Pallecchi, L. et al., 2007. Population structure and resistance genes in antibiotic-resistant bacteria from a remote community with minimal antibiotic exposure. *Antimicrobial Agents and Chemotherapy*, 51(4), pp.1179–1184.

Pinheiro, J. et al., 2007. Linear and nonlinear mixed effects models. R package version. Available at: http://CRAN.R-project.org/package=nlme.

R Core Team, 2013. R: A language and environment for statistical computing. Available at: http://www.R-project.org.

Sawant, A.A. et al., 2007. Antimicrobial-resistant enteric bacteria from dairy cattle. *Applied and Environmental Microbiology*, 73(1), pp.156–163.

Schmid-Hempel, P., 2001. On the evolutionary ecology of host-parasite interactions: addressing the question with regard to bumblebees and their parasites. *Naturwissenschaften*, 88(4), pp.147–158.

Slater, F.R. et al., 2008. Progress towards understanding the fate of plasmids in bacterial communities. *FEMS Microbiology Ecology*, 66(1), pp.3–13.

Smalla, K. & Sobecky, P.A., 2002. The prevalence and diversity of mobile genetic elements in bacterial communities of different environmental habitats: insights gained from different methodological approaches. *FEMS Microbiology Ecology*, 42(2), pp.165–175.

Smalla, K., Osborn, A.M. & Wellington, E.M.H., 2005. Isolation and characterisation of plasmids from bacteria. In C. M. Thomas, ed. *The horizontal gene pool: bacterial plasmids and gene spread*. Amsterdam: Taylor and Francis, pp. 212–251.

Sobecky, P.A., 2002. Approaches to investigating the ecology of plasmids in marine bacterial communities. *Plasmid*, 48(3), pp.213–221.

Sobecky, P.A. et al., 1997. Plasmids isolated from marine sediment microbial communities contain replication and incompatibility regions unrelated to those of known plasmid groups. *Applied and Environmental Microbiology*, 63(3), pp.888–895.

Spellerberg, I.F. & Fedor, P.J., 2003. A tribute to Claude Shannon (1916-2001) and a plea for more rigorous use of species richness, species diversity and the 'Shannon-Wiener' Index. *Global Ecology and Biogeography*, 12(3), pp.177–179.

Thomas, C.M., 2004. Evolution and Population Genetics of Bacterial Plasmids. In B. E. Funnell & G. J. Phillips, eds. *Plasmid Biology*. Washington: American Society for Microbiology (ASM), pp. 509–528.

Top, E. et al., 1994. Exogenous isolation of mobilizing plasmids from polluted soils and sludges. *Applied and Environmental Microbiology*, 60(3), pp.831–839.

Van den Broeck, F. et al., 2014. Inbreeding within human *Schistosoma mansoni*: do host-specific factors shape the genetic composition of parasite populations? *Heredity*, 113(1), pp.32–41.

Van Kessel, J. et al., 2007. Survival of *Escherichia coli* in cowpats in pasture and in laboratory conditions. *Journal of Applied Microbiology*, 103(4), pp.1122–1127.

Villa, L. et al., 2010. Replicon sequence typing of IncF plasmids carrying virulence and resistance determinants. *Journal of Antimicrobial Chemotherapy*, 65(12), pp.2518–2529.

Wang, L. et al., 2003. Species-wide variation in the *Escherichia coli* flagellin (H-antigen) gene. *Journal of Bacteriology*, 185(9), pp.2936–2943.

Wang, Y. et al., 2006. Antibiotic resistance and plasmid profile of environmental isolates of *Vibrio* species from Mai Po Nature Reserve, Hong Kong. *Ecotoxicology (London, England)*, 15(4), pp.371–378.

Wright, S., 1978. Variability within and among natural populations. In *Evolution and the genetics of populations : a treatise in four volumes*. University of Chicago Press, pp. 246–326.

Chapter 6: Parasitic plasmid persistence and conjugation

6.1 Introduction

6.1.1 Plasmid persistence and parasitism

Three fundamental forces governing plasmid persistence were identified by Levin and Stewart over 30 years ago: the rate of conjugation, the rate of segregational loss and the cost of plasmid carriage to the host (Stewart & Levin 1977; Levin & Stewart 1980). For plasmids to persist in a population, the relationship between these parameters can be summarized by the equation:

Transfer Rate > Cost of Carriage + Segregation Rate

Plasmids can be lost from a population by segregation, where plasmids are not equally distributed amongst daughter cells during cell division. This segregational loss may be exacerbated by a high cost of plasmid carriage: more costly plasmids may be lost at a greater rate. Conversely, these factors may be mitigated by high rates of conjugation: horizontal transfer to plasmid-free cells.

The survival of plasmids which provide some selective advantage to their host in a specific environment appears to be straight forward (Stewart & Levin 1977), and persistence through the carriage of transiently beneficial genes, which facilitate growth in a particular niche (an 'accessory genome'), is a common explanation (Eberhard 2003). However, although this theory could explain the localization of many apparently non-essential genes to plasmids, and their lack of essential 'housekeeping' genes (although see section 6.1.3 below, and Galibert et al. (2001), it invokes the controversial theory of population-level selection, and fails to explain the existence of plasmids carrying beneficial traits in the face of transfer of these genes to the chromosome (Thomas 2004; Eberhard 1990; Bergstrom et al. 2000). As Eberhard (1990) bluntly states *"this and other explanations… are probably wrong*". In reality, not all plasmids confer an obvious advantage, and beneficial plasmids are also found beyond their adaptive niche.

The challenge is to explain the survival of these cryptic and residual plasmids. Theory predicts that plasmids can be maintained parasitically, without conferring an advantage to their host if the parameters of the above equation can be balanced appropriately, i.e. high conjugation, low cost and low segregation rate (Figure 6.1). However, this has been difficult to demonstrate experimentally (Gordon 1992; Bahl et al. 2007; Freter et al. 1983; Simonsen 1991).

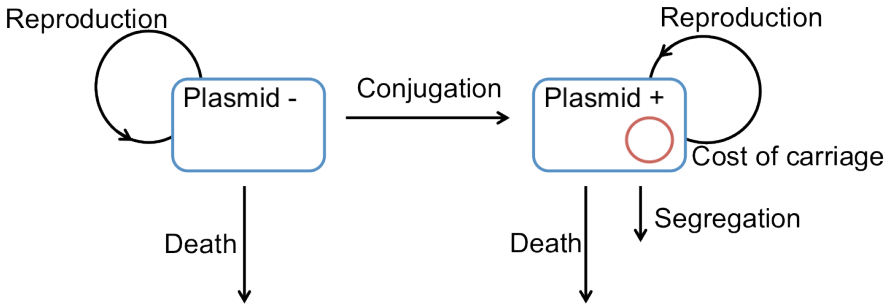

Figure 6.1: Factors affecting the survival of plasmid bearing and plasmid free cells. The plasmid bearing cells must balance the cost of plasmid carriage and the segregation rate with conjugation rate in order to survive. Adapted from Lili et al. (2007).

The rate of conjugation is key to the argument against parasitism: are rates of transfer high enough in 'real' populations for plasmids to be maintained parasitically? Early work by Lundquist & Levin (1986) demonstrated the importance of conjugation for plasmid carrier invasion from rare in laboratory broth culture, but only two of the seven plasmids tested were able to invade. In addition, both this study and the subsequent work by Simonsen et al (1990) used laboratory plasmids almost exclusively. It is likely that transfer rates of wild plasmids would differ from those produced artificially in the laboratory. Indeed transfer rates of plasmids isolated from wild *E. coli* populations have been shown to be too low to support parasitic plasmid maintenance based on Stewart & Levin's models (Gordon 1992).

6.1.2 Alternative models of plasmid persistence

The 'local adaptation hypothesis', proposed by Eberhard (1990), states that new species colonizing a site may obtain beneficial site-specific genes by plasmid acquisition, and that in return, plasmid transfer to incoming clones will result in plasmid association with superior (fitter) host backgrounds, allowing the plasmid to increase in frequency. Bergstrom et al. (2000) expand upon these ideas and update the 1977 model (Stewart & Levin 1977) to accommodate reportedly low conjugation rates (Gordon 1992) and the transfer of beneficial genes to the chromosome. They propose that over time, plasmid carried beneficial genes will move to the chromosome and plasmids will be lost unless there are regular selective sweeps through the host population. Plasmids may transfer to the incoming (fitter) host background and thus 'hitchhike' to success. Plasmid carried gene copies may also transfer across species and strain boundaries, thus avoiding deletion through selective sweeps in one host. These scenarios both emphasize the importance of conjugation but struggle to explain plasmid persistence in the absence of a benefit to the host, suggesting that cryptic plasmids either: i) have a hidden benefit we are unaware of; ii) are 'in transit' and must gain a beneficial gene or be lost; or iii) exist parasitically by evolving to have almost no cost.

Reevaluation of the models derived from Stewart & Levin (1977) has led to a very different conclusion: that plasmids can persist parasitically via oscillations between plasmid-bearing and plasmid-free cells, negating the need for complex heterogeneous population models such as hitchhiking and ecotype switching, and allowing stable existence of cryptic plasmids (Lili et al. 2007). This stable coexistence of plasmid-free and plasmid-bearing cells is also supported by a novel modeling approach allowing the burden of plasmid carriage to change stochastically, independently of previous time points (Ponciano et al. 2007). This method is biologically neat, as it allows environmental and ecological factors affecting plasmid cost to be modeled, so that for example, compensatory mutations could reduce the cost of carriage or that removal of a selective pressure (i.e. antibiotic treatment) could increase the cost. It should be noted that the effect of coevolution of a host/plasmid combination is likely to lead to a

downward trend in plasmid cost, rather than stochastic variation (Dionisio et al. 2005; Dahlberg & Chao 2003; Bouma & Lenski 1988). Moreover, the model explains only some of their data (four out of seven experiments) better than models derived from Stewart & Levin, (1977). However, as with the model proposed by Lili et al. (2000), this model predicts some stable coexistence of plasmid-free and plasmid-bearing strains, and additionally allows estimation of the key parameter causing plasmid loss: segregation rate, plasmid cost or conjugation rate (Figure 6.1).

6.1.3 Plasmid and host (co-) evolution

Evolution of the plasmid, the host cell and/or coevolution of the two can have a major impact on the key parameters of plasmid survival (Figure 6.1). Plasmid evolution can reduce the cost of carriage to zero over just 420 generations (Dionisio et al. 2005). In contrast, host evolution over 500 generations (with no detectable change in the plasmid) can not only reduce to the cost of carriage to zero, but can also confer a significant fitness advantage to plasmid carriers (Bouma & Lenski 1988). There is also middle way: coevolution of the host and plasmid (Dahlberg & Chao 2003). Harrison & Brockhurst (2012) propose three mechanisms by which coevolution can reduce the cost of plasmid carriage: altered conjugation rate, loss of genes and changes in gene expression.

The first mechanism they propose is altered conjugation rate. This may seem inexplicable at the gene or plasmid level, but could be beneficial at the level of the cell: conjugation is metabolically costly, inhibits (temporarily) replication, and may make cells more vulnerable to phage attack (Thomas 2004; Eberhard 1990). Plasmid evolution can be seen as analogous to the evolution of symbiosis, where coadaptation of plasmid and host could be a step on the road to obligate symbiosis and the 'domestication' of plasmids (Slater et al. 2008). Slater et al. cite the example of the large, apparently obligate-symbiont plasmids of *Sinorhizobium meliloti.* One of these plasmids, pSymB, appears to have lost most of its transfer genes, and gained an essential housekeeping gene, suggesting that it is well down the path towards obligate symbiosis (Galibert et al. 2001). The

host can also affect the transfer rate by modulating plasmid gene expression, highlighting the importance of coevolution of host and plasmid (Dahlberg & Chao 2003). The second mechanism proposed is loss of genes. This seems a logical way to reduce the cost of plasmid carriage, as fewer genes means less DNA to replicate, as well as reducing the cost of gene expression. However, it seems that this mechanism is less common in reality than might be expected. Dahlberg & Chao (2003) examined the evolution of two plasmids and their hosts over 1100 generations in antibiotic-free broth culture, and found that most lineages retained a costly antibiotic resistance gene. In addition, just one replicate had reduced plasmid copy number - another form of gene loss - and moreover, one replicate had increased copy number. Again, this appears to be a complex mechanism, perhaps due to the tension between copy number and plasmid stability (Thomas 2004). The final mechanism, plasmid cost, may be reduced by altered gene expression, regulated by either the plasmid or the host. As mentioned above, Dahlberg & Chao suggest that conjugation is limited in their evolved strains by actions of the host to suppress gene expression. However, tight regulation of gene expression can also be encoded by the plasmid, as seen in the Inc-P1 family (Schlüter et al. 2007; Heuer et al. 2007).

Although the key factors affecting plasmid maintenance remain true, subsequent experimental and theoretical work has suggested that this is perhaps a simplification, and that other factors such as population density, environmental fluctuation and (co-) evolution of host and plasmid are equally important (Thomas 2004; Harrison & Brockhurst 2012; Slater et al. 2008; Simonsen 1991). Furthermore, it seems that estimations of these key parameters affecting plasmid survival are still open to debate. It is nevertheless clear that conjugation plays a critical role in plasmid persistence (Sia et al. 1995; Easter et al. 1997), and that horizontal transfer does occur in biologically meaningful contexts, such as the gut (Freter et al. 1983).

6.1.4 Aims

This chapter aims to investigate the importance of conjugation for plasmid persistence, and to explore its potential to facilitate 'hitchhiking' to incoming strains, as proposed by Bergstrom et al. (2000). This work focuses on the maintenance of plasmids in a non-selective environment, as parasitic, 'residual' elements. The plasmid used, pCT, is a naturally occurring resistance plasmid (Cottell et al. 2011), transplanted into a non-selective environment and a non-coevolved host. Studies in this system facilitate investigation of how large plasmids may survive once they are no longer under selection, and how evolution may lead to extinction or some other fate such as becoming a cryptic plasmid. As well as giving an insight into mechanisms of plasmid persistence, this is key to understanding the spread of resistance plasmids in clinical and community settings.

Serial transfer experiments using the conjugative plasmid, pCT, and its non-conjugative derivative, pCT Δ*trbA,* with plasmid-free competitors, were used to test the hypothesis that in a non-selective environment plasmid-carrying cells would be outcompeted by their plasmid-free counterparts, and that the decline would be accelerated for non-conjugative plasmids. These experiments also tested the effect of bottleneck intensity on the rate of plasmid-carrier decline. Secondly, serial transfer experiments were used to test the hypothesis that the proportion of plasmid carriers will increase faster if an incoming recipient strain is fitter than the current plasmid host. The results of these experiments were surprising, and necessitated further experimental work to ascertain the reasons for this, including examination of evolved strains for altered fitness or conjugational ability, and testing for frequency-dependent fitness effects.

6.2 Methods

6.2.1 Strains & culture techniques

The strains used in this chapter are *E. coli* K-12 MG1655 and the MG1655 Δ*lacZYA* mutant. The plasmids are pCT and its non-conjugative derivative pCT Δ*trbA.* pCT carries the ESBL resistance gene CTX-M-14 (Cottell et al. 2011). Plasmids were transformed into both host backgrounds. The wild type MG1655 strain is blue in the presence of X-Gal & IPTG whereas the mutant Δ*lacZYA* strain is white. The fitness cost of the *lacZYA* knockout was assessed with competition in broth, the results of which can be seen in Chapter 2, Section 2.3.3. The blue MG1655 strain appears to be slightly fitter in broth. Generic culture techniques and antibiotic specifications can also be seen in Chapter 2, Section 2.2.1.

6.2.2 Serial transfer experiments

Serial transfer experiments were conducted in 1 ml LB broth + X-Gal (0.02 mg ml^{-1}) & IPTG (0.1 mM) (Fisher Scientific) in 24 well plates (Costar® Corning®, Sigma-Aldrich). All wells contained broth, but only alternate wells were inoculated to reduce and monitor for cross-contamination (Figure 6.2). Experimental mixes were initiated from fresh overnight cultures of two strains: four combinations of plasmid carriers and plasmid-free cells were used (see further details below and Table 6.1). OD_{600} of these overnight cultures was measured and strains were diluted to approximately equal density, and then mixed to give the desired initial proportion of plasmid carriers. Inoculated 24 well plates were incubated at 37°C and shaken at 120 rpm. Cultures were transferred to fresh broth every 2-3 days in a Monday, Wednesday, Friday pattern. All transfers were conducted using filter tips in a laminar flow cabinet to reduce contamination risk, as no selective antibiotic was used.

Samples were taken periodically from the cultures before transfer in order to measure the proportion of plasmid carrying cells. These samples were serially diluted, plated onto LB Agar + X-Gal & IPTG with and without cefotaxime and

incubated at 37°C overnight. This allowed counts of plasmid-free cells and plasmid carriers, including transconjugants, to be obtained. Estimates of the number of segregants were made by subtracting counts of original plasmid carriers on cefotaxime from counts on LB only plates. Samples from the serial transfers were also periodically stored in 80% glycerol at -80°C (Chapter 2, Section 2.2.1).

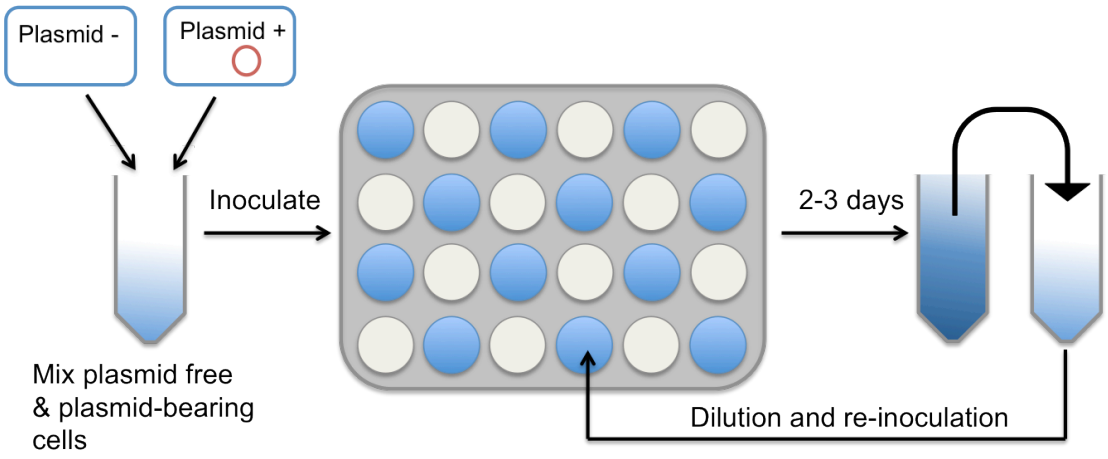

Figure 6.2: Serial transfer experiments. Mixtures of plasmid-free and plasmid-bearing cells were inoculated into alternate wells of a 24 well plate, containing LB broth + X-Gal & IPTG. Cultures were incubated at 37°C and shaken at 120 rpm. Transfers occurred every 2-3 days, when cultures were diluted and re-inoculated into fresh broth. Samples were taken periodically before transfer and plated onto cefotaxime containing plates to determine the proportion of plasmid-carrying cells.

Serial transfer experiment 1

Experiment 1 used MG1655 + pCT and MG1655 + pCT Δ*trbA* as the plasmid carrying strains, with MG1655 Δ*lacZYA* as the plasmid-free strain in both cases, and an initial proportion of plasmid carriers of approximately 0.01. Three transfer regimes were used in this experiment to vary the bottleneck intensity. At transfer, the old culture was diluted with 0.85% (w/v) saline either 1:10, 1:100 or 1:1000, and fresh broth was inoculated by a further 1:10 dilution. The total dilutions at transfer (including both saline dilution and broth inoculation)

were therefore 100-, 1000- and 10000-fold for the low, medium and high bottleneck intensities respectively. There were six replicates for each treatment and the experiment lasted 21 days.

Serial transfer experiment 2

Experiment 2 used four combinations of plasmid-free and plasmid carrying cells, with both plasmids in both host backgrounds (Table 6.1). Two initial proportions of plasmid carriers were used for each mixture, approximately 0.0001- 0.00001. The two strains were added to the broth separately, rather than producing an inoculation mixture. All transfers were conducted with moderate bottleneck intensity (1000-fold total dilution). There were six replicates of each treatment, and 2 independent replicates of the whole experiment. For simplicity data is only presented here for the longer of these two experimental replicates, lasting 32 days compared to 17 days. However, both experimental replicates show the same pattern over the first 17 days.

Blue strain	**White strain**
MG1655 + pCT	MG1655 Δ*lacZYA*
MG1655 + pCT Δ*trbA*	MG1655 Δ*lacZYA*
MG1655	MG1655 Δ*lacZYA* + pCT
MG1655	MG1655 Δ*lacZYA* + pCT Δ*trbA*

Table 6.1: Combinations of plasmid carriers and plasmid-free cells. All four combinations were used in serial transfer 2, only the first two (with the plasmids in the MG1655 host) were used in serial transfer 1.

6.2.3 Serial transfer 2 end point assays

Selected strains from serial transfer 2, where the initial proportion of plasmid carriers was varied, were assessed for evolution of the host and plasmid after serial transfer for 32 days. Conjugative pCT carrying strains were selected to assess any change in the cost of plasmid carriage or rates of conjugation. In

addition, a strain that maintained a particularly high proportion of plasmid carriers in the serial transfer (MG1655 *ΔlacZYA* + pCT *ΔtrbA* evolved strain 2B1), was assessed for altered fitness in comparison to the ancestor. Strains were revived from serial transfer glycerol stocks and streaked out twice to obtain pure single colonies.

6.2.3.1 Competition experiments

Competition experiments were conducted to assess the relative fitness of the evolved and ancestral lineages, as described in Chapter 2, Section 2.2.4. Evolved lineages were competed against their opposite ancestral type, for example evolved MG1655+pCT would compete against ancestral MG1655 Δ*lacZYA*. Competitions were conducted in 1 ml LB broth + X-Gal & IPTG in 24 well plates. Overnight cultures of the competitors were diluted to approximately equally cell density and mixed 1:1, then inoculated into broth by a 1:5 dilution. 20 µl was immediately taken and diluted to obtain the starting ratio. Broth cultures were incubated at 37°C and shaken at 120 rpm for 24 hours. Final counts were obtained by serial dilution and plating onto LB agar + X-Gal & IPTG +/- cefotaxime. Relative fitness in broth competitions was calculated as described by Ross-Gillespie et al. (2007):

$$v = x_2(1 - x_1)/ x_1(1 - x_2)$$

Where v is the relative fitness of MG1655, x_1 is their initial proportion and x_2 their final proportion.

These competition experiments were conducted for the four conjugative pCT carrying strains, two with pCT white background (1A3, 1B6) and two in blue background (3A6, 3B6). Six replicates, initiated from independent colonies, were performed for each strain. For strain 1A3, 2 original plasmid carriers (white) and 4 transconjugants (blue) were used. Competitions were also conducted with the evolved strain 2B1 and its ancestor MG1655 *ΔlacZYA* + pCT *ΔtrbA.* Twelve replicates, initiated from independent colonies, were used in these two competitions.

6.2.3.2 Mating experiments

Mating experiments were conducted to assess any changes in conjugation rate between evolved and ancestral lineages. Crosses were set up from independent donor and recipient cultures, and evolved lineages were again mated with their opposite ancestral type. OD_{600} of overnight cultures were measured and cultures diluted to approximately the same density. Mixtures were made with a 1:5 ratio of donors to recipients and a 1 in 5 dilution into fresh LB broth + X-Gal & IPTG, with a final volume of 1 ml. Mating experiments were conducted in 24 well plates, and incubated at 37°C, 220 rpm for 24 hours. Initial and final mixtures were serially diluted and plated out onto LB + X-Gal & IPTG +/- cefotaxime to assess ratios. Conjugation rate was calculated by the equation:

$$Transconjugants\ per\ donor = \frac{Total\ transconjugants}{Total\ donors}$$

6.2.4 Frequency dependence assay

Data from serial transfer experiments indicated that plasmid carrier fitness might have a frequency dependent element. In order to assess this a frequency dependence assay was conducted. Here, competition experiments were carried out as described above (Section 6.2.3.1) over a range of initial proportions of plasmid carriers from 0.01-0.9, using the non-conjugative pCT Δ*trbA* plasmid in both host backgrounds.

6.2.5 Statistical analysis

The serial transfer experiments were designed such that a linear mixed-effects (lme) model could be used to analyze the resulting data, with the proportion of plasmid carriers as the response variable; time, plasmid type, donor type and transfer regime as fixed effects; and replicate as a random effect. Modeling was conducted in *R* (R Core Team 2013) using the *nlme* package (Pinheiro et al. 2007). Restricted maximum likelihood (REML) lme models were converted to maximum likelihood (ML) models to allow model comparison using mixed model ANOVA. End point assays were compared using ANOVA or t tests as appropriate.

6.3 Results

6.3.1 Plasmid carriers decline from relatively rare

In serial transfer 1, the intensity of the transfer bottleneck was varied from low (100-fold dilution) to moderate (1000-fold) or high (10000-fold). Under all bottleneck regimes the proportion of plasmid carriers declines, regardless of the conjugational ability of the plasmid. In addition, the rate of decline is more rapid for both plasmids at higher bottleneck intensities (Figure 6.3). This can be seen in the significant interaction effect between dilution (1/bottleneck regime) and time in the model (p = 0.001, Table 6.2). For all treatments there is an approximate 10-fold decline in the proportion of plasmid carriers over the first seven days, from the initial proportion of 0.01 (data not shown). Subsequently the non-conjugative plasmid, pCT Δ*trbA,* shows a steeper decline in plasmid carrying cells under all bottleneck regimes than its conjugative precursor, as demonstrated by the significant interaction effect between plasmid type and time in the model (p = 0.028, Table 6.2). Non-conjugative plasmid carriers drop to undetectable levels (< 1 CFU in 20 μl undiluted culture) in all six replicates in the moderate regime and four out of six in the high-intensity regime after just 14 days (Figure 6.3 B & C). By 21 days, plasmid carriers were undetectable in all replicates in the high and moderate intensity regimes, and three out of five replicates in the low intensity regime, with the other two showing a considerable reduction (Figure 6.3 A). The proportion of conjugative plasmid carriers also declined in almost all replicates across the three regimes. For the high intensity regime, the results are very similar to the non-conjugative plasmid; four out of six replicates had no detectable plasmid carriage after 14 days, and by 21 days plasmid carriers were undetectable in all six treatments. A less dramatic decline is seen in the moderate intensity regime with half of the replicates maintaining detectable levels of plasmid carriage at 21 days. The low intensity regime shows the most marked difference from the non-conjugative plasmid, with just one in six having undetectable resistance after 21 days. In addition to this, one replicate actually maintained a relatively stable proportion of pCT carriers; after an initial decline from 0.01 to 0.00037 at 7 days, the proportion rose to 0.0019 at 14 days and 0.0038 at 21 days (Figure 6.3 A).

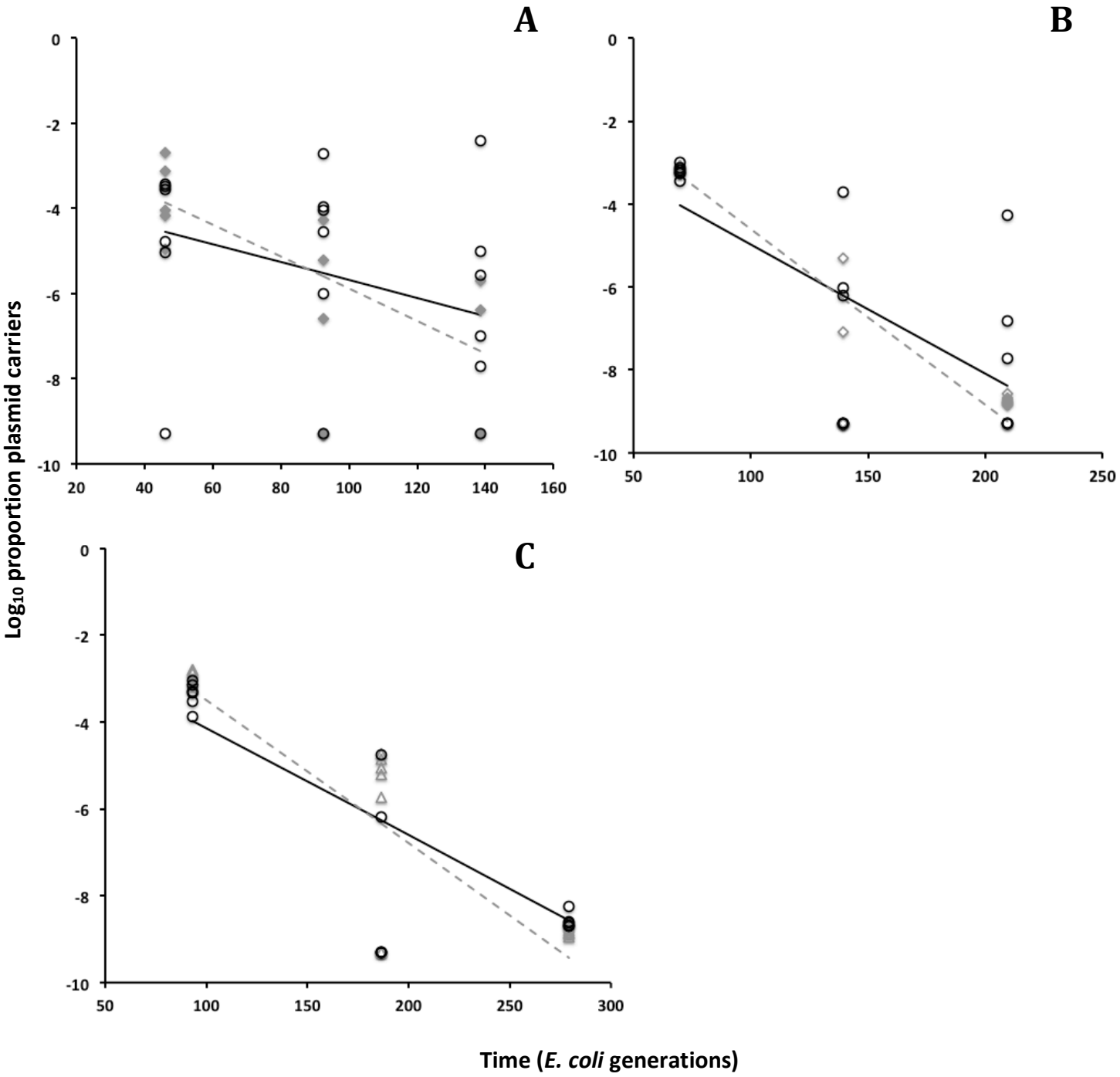

Figure 6.3: Plasmid carriers decline in mixed non-selective serial transfer. The rate of decline is faster for non-conjugative plasmids, and decreases with bottleneck intensity from low **(A)** to moderate **(B)** and high **(C)**. The mean rate of decline is significantly slower for wild type pCT (circles & solid line) carriers than non-conjugative pCT Δ*trbA* (diamonds & broken line) ($p = 0.028$), and under low bottleneck intensity (A) one replicate retains pCT carriage. There is a significant interaction between the bottleneck regime and time ($p = 0.001$), indicating that plasmid carriers decline faster at higher bottleneck intensity. The plasmid host background was MG1655 and its competitor was plasmid-free MG1655 Δ*lacZYA.* Serial transfer was in non-selective LB broth + X-Gal & IPTG, with transfers every 2-3 days. Lines connect mean predicted values from the model (Table 6.2).

Model	Term removed	AIC	ΔAIC	*p*
Maximal model	-	414.415	0	-
2.	plasmid: time: dilute	415.076	+0.661	0.103
Minimal adequate model	plasmid: dilute	413.730	-1.346	0.418
4.	plasmid: time	416.551	+2.821	0.028*
5.	dilute: time	422.110	+8.380	0.001*
Null model	Mean only	500.432	+86.702	<0.0001*

Table 6.2: Model simplification for serial transfer 1. Serial transfer data was modeled using a linear mixed effect model with the proportion of plasmid carriers as the response variable, replicate as the random effect and time, plasmid type and dilution as fixed effects. The proportion plasmid carriers was $\log_{10}$ transformed after addition of the minimal detectable value (5×10^{-10}). "Dilute" represents 1/ bottleneck regime i.e. for the moderate intensity regime: 1/1000 = 0.001. ":" indicates an interaction term. Maximal model:

Proportion plasmid carriers ~ plasmid: time: dilute + plasmid: dilute + plasmid: time + dilute: time + plasmid + time + dilute, random effect = replicate.

Terms were removed sequentially from the maximal model, and models compared using ANOVA. Simplifications were accepted if the models were not significantly different when a term was removed (*p* * indicates a significant difference). AIC (Akaike's information criterion) values were also compared for a general downward trend. Note that the AIC values of model 5 and the null model are compared to the minimal adequate model rather than the previous iteration, as models 4 and 5 excluded significant terms. Main effects were not tested separately due to their inclusion in interaction terms (Nelder 1977). The null model is also given for comparison. The minimal adequate model has significant effects of the plasmid: time interaction and the dilute: time interaction, as given by the model:

Proportion plasmid carriers ~ plasmid: time + dilute: time + plasmid + time + dilute, random effect = replicate.

Although Figure 6.3 suggests that conjugation may contribute somewhat to plasmid maintenance, Figure 6.4 shows that the effect of conjugation is highly variable between conjugative plasmid replicates. In three cases (replicates 3 and 4 of the low intensity regime and replicate 6 in the moderate regime) plasmid persistence is entirely dependent on transconjugants. However, in one replicate of the low intensity regime (where pCT is maintained at a relatively stable proportion) and one replicate in the moderate regime no transconjugants were detected at all.

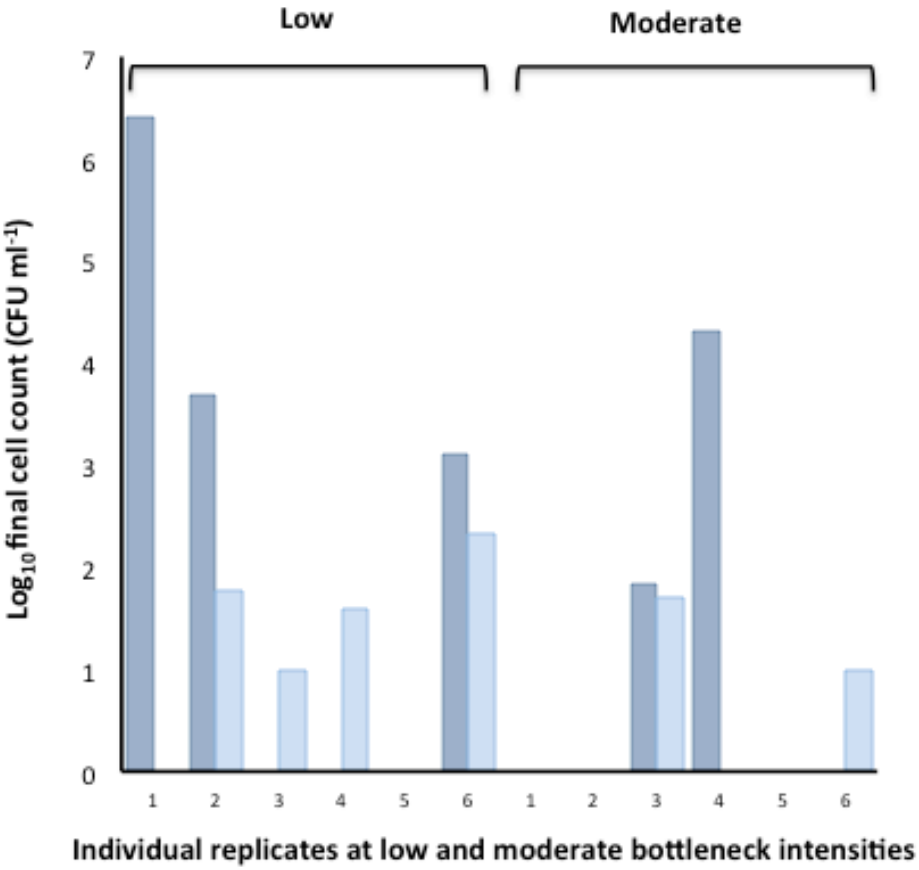

Figure 6.4: The contribution of original plasmid carrying cells and transconjugants to final plasmid carrying cell counts. Original plasmid carrying cells (dark blue) and transconjugants (light blue) (log_{10} CFU ml^{-1}) in the low and moderate bottleneck intensity regimes. Bars represent cell counts at day 21 for individual serial transfer replicates with the conjugative pCT plasmid. Data is not shown for the high intensity regime, as plasmid carriers were undetectable in all replicates.

It is possible that plasmid carriers were maintained by conjugation with segregants that had lost the plasmid. An attempt was made to estimate numbers of segregants in serial transfer 2 by subtracting the number of plasmid carriers on ampicillin plates from the number on antibiotic free plates. There are limitations to this method, as it is difficult to identify plasmid carriers on non-selective plates amongst high densities of plasmid free cells, and in fact this estimate suggests wildly variable segregation rates, with the proportion of original plasmid carriers that are segregants ranging from 0 to 0.99 (for pCT Δ*trbA* in the white background at day 5). Although it is not possible to draw any conclusions about the importance of conjugation with segregants, it is clear that transconjugants do contribute to plasmid survival in both the low and moderate intensity bottleneck regimes.

Across all three bottleneck regimes it can be seen that conjugation is not sufficient to maintain plasmid carriers, as levels drop below detection in 1 of 6, 3 of 6, and 6 of 6 cases for the low, moderate and high intensity regimes respectively. However, the conjugative plasmid does have slightly lower rate of decline than its non-conjugative derivative (p = 0.028, Table 6.2), and the decline is slower with less intense bottlenecks (Figure 6.3, p = 0.001). This data highlights the importance of population dynamics for plasmid maintenance: with relatively low initial plasmid frequency (around 0.01) and frequent high-intensity bottlenecks, conjugation cannot maintain the plasmid in the absence of selection, while conjugation may be able to maintain the plasmid if the bottleneck intensity is low.

6.3.2 Plasmid carriers increase from extremely rare

It can be seen from Figure 6.2 that there is a marked difference in plasmid survival between low and moderate intensity bottlenecks. In order to investigate whether hitchhiking can maintain the plasmid in the face of moderate intensity bottlenecks (Bergstrom et al. 2000), the moderate intensity serial transfer was replicated with plasmids in both the white (MG1655 Δ*lacZYA*) and blue (MG1655) backgrounds. The results from Chapter 2 indicate that the blue strain

has slightly higher fitness in broth (Chapter 2, Figure 2.5), and therefore it is predicted that the plasmid should hitchhike to higher frequencies when moving from the white to the blue background. In order to force rapid extinction of the non-conjugative plasmid, serial transfers were initiated with low proportions of resistant cells, from 0.001- 0.00001 (Figure 6.5). The results are surprising. In all cases, the proportion of resistant cells increases to 0.01 – 0.1. The fact that all plasmids increase in frequency, regardless of conjugational ability or host background, suggests that neither conjugation nor hitchhiking are key factors driving the increase.

Although plasmid type has a significant effect on the proportion of plasmid carriers over time (p = 0.009, Table 6.3), the pattern for both plasmids is very similar over the first 19 days. This can be seen by the parallel curves of the model (Figure 6.5). The non-conjugative plasmids had a slightly higher initial proportion, but otherwise the shapes of the curves are the same. After 19 days the non-conjugative plasmid actually declines slightly less steeply than conjugative pCT, suggesting that conjugation is not a key contributory factor to plasmid survival, despite the fact that transconjugants account for a large proportion of plasmid carriers in the pCT mixtures (Figure 6.6). This difference in final proportion of plasmid carriers is curious, and could perhaps be attributed to reduced fitness cost of pCT Δ*trbA* carriage due to the deletion, although it is also possible that this difference is merely a fluctuation that would even out over time.

Contrary to the hitchhiking hypothesis, the model shows that plasmid host background ("donor") does not have a significant effect on the proportion of plasmid carriers over time, and was in fact excluded from the model (p = 0.125, Table 6.3). This confirms that hitchhiking is not a contributory factor for plasmid prevalence in this experiment. This also corroborates the suggestion that conjugation is not a significant factor for plasmid survival as the non-conjugative plasmid does just as well, if not better, than its conjugative counterpart (Figure 6.5).

A second non-linear model with temporal pseudoreplication was used to incorporate an asymptotic relationship between plasmid frequency and time (Figure 6.5C). The AIC value for this model is 816.81, almost identical to the value for the linear model (AIC = 816.87, Figure 6.5 A & B), indicating that the models have similar explanatory power. However this model is more biologically appropriate as the asymptotic shaped curve is often associated with negative frequency dependent selection (see Figure 6.9 below).

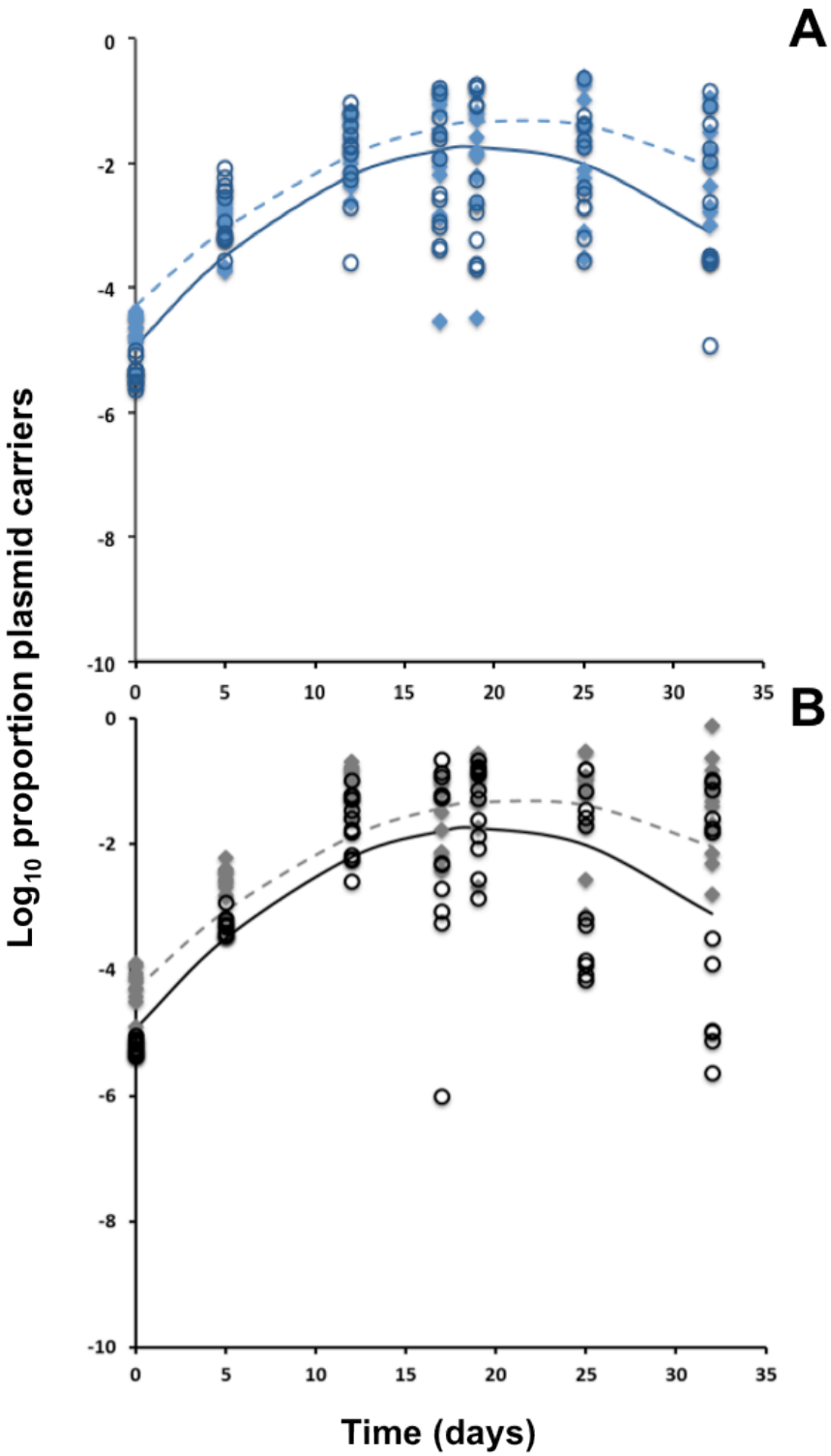

Figure 6.5: The proportion of plasmid carriers can increase from rare and can be maintained in the absence of selection. (A) plasmids in blue MG1655 background in competition with MG1655 Δ*lacZYA;* **(B)** plasmids in white MG1655 Δ*lacZYA* in competition with MG1655. Dark-coloured circles & solid line: wild type pCT. Light-coloured diamonds & broken line: non-conjugative pCT Δ*trbA*. Lines connect mean predicted values from the model (Table 6.3). Serial transfer was in non-selective LB broth + X-Gal & IPTG, with moderate intensity bottlenecks and transfers every 2-3 days. There are approximately 10 *E. coli* generations per day in this experiment.

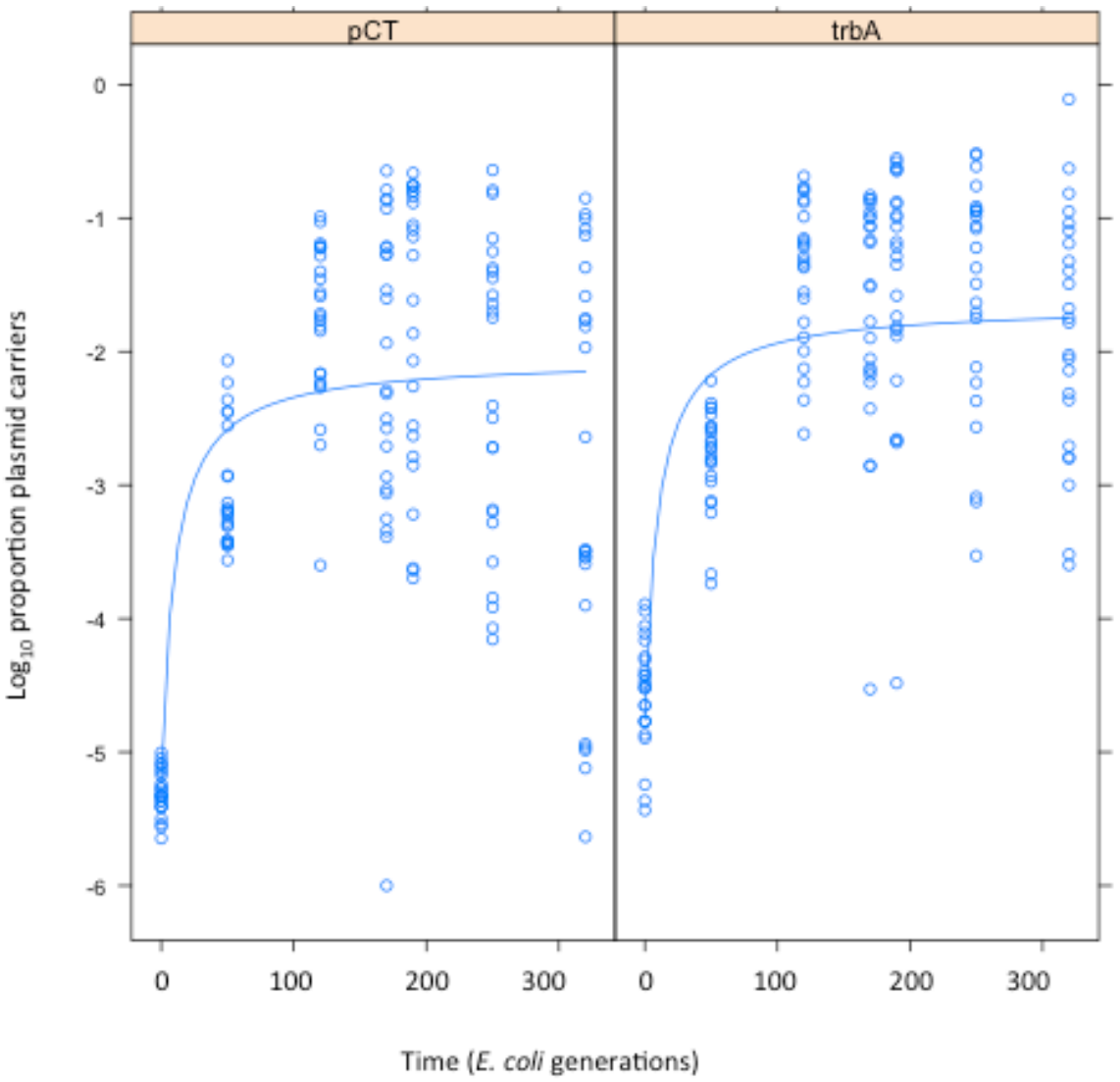

Figure 6.5 C: Asymptotic model of proportion of plasmid carriers against time. 'Donor' was shown to be non-significant in the linear model, and was therefore not included in this model. Replicates from both donor backgrounds are pooled for the pCT plasmid (left panel) and pCT Δ*trbA* plasmid (right panel). A 2-parameter asymptotic exponential function of time was fitted $y = a\ (1\text{-}\ e^{\text{-}bx})$, where a and b were estimated from the data (a = -2; b = 3, AIC = 816.81). A three-parameter model, $y = a\ \text{-}\ be^{\text{-}cx}$, was not significantly different from the two parameter model (ANOVA, p = 0.366, AIC = 820.50). The model formula is:

$$\text{Proportion plasmid carriers} \sim a\left(1 - e^{-b*\text{time}}\right)|\ \text{plasmid}$$

Model	Term removed	AIC	ΔAIC	*p*
Maximal model	-	818.754	0	
2.	plasmid: donor: time	816.459	-2.295	0.774
3.	plasmid: donor: time^2	817.213	+0.754	0.123
4.	donor: time^2	817.059	-0.154	0.174
5.	donor: time	815.466	-1.593	0.523
6.	plasmid: donor	816.522	+1.056	0.081
7.	plasmid: time	818.488	+1.966	0.046*
8.	plasmid: time^2	821.266	+4.744	0.009*
Minimal adequate model	donor	816.872	+0.35	0.125
Null model	Mean only	1184.158	+367.286	<0.0001*

Table 6.3: Model simplification for serial transfer 2. Serial transfer data was modeled using a linear mixed effect model with the proportion of plasmid carriers as the response variable, replicate as the random effect and time, time^2, plasmid type and donor type as fixed effects. Donor type indicates the initial host background of the plasmid (blue or white). The proportion of plasmid carriers was $\log_{10}$ transformed after addition of the minimal detectable value (1×10^{-6}). The quadratic term, time^2, is included in the model to allow for the curvature of the data, but interaction terms including time^2: time were excluded, as this produces the cubic term, time^3, creating an inappropriate curve. ":" indicates an interaction term. Maximal model:

Proportion plasmid carriers ~ plasmid: donor: time + plasmid: donor: time^2 + donor: time^2 + donor: time + plasmid: donor + plasmid: time + plasmid: time^2 + plasmid + donor + time + time^2, random effect = replicate.

Terms were removed sequentially from the maximal model, and models compared using ANOVA. Simplifications were accepted if the models were not significantly different when a term was removed (*p* * indicates a significant difference). AIC (Akaike's information criterion) values were also compared for a general downward trend. Note that the AIC values of model 8 and the minimal adequate model are compared to model 6, rather than the previous iteration, as models 7 and 8 excluded significant terms. Main effects were not tested separately when included in interaction terms (Nelder 1977). The null model is also given for comparison. The minimal adequate model is:

Proportion plasmid carriers ~ plasmid: time + plasmid: time^2 + plasmid + time + time^2, random effect = replicate.

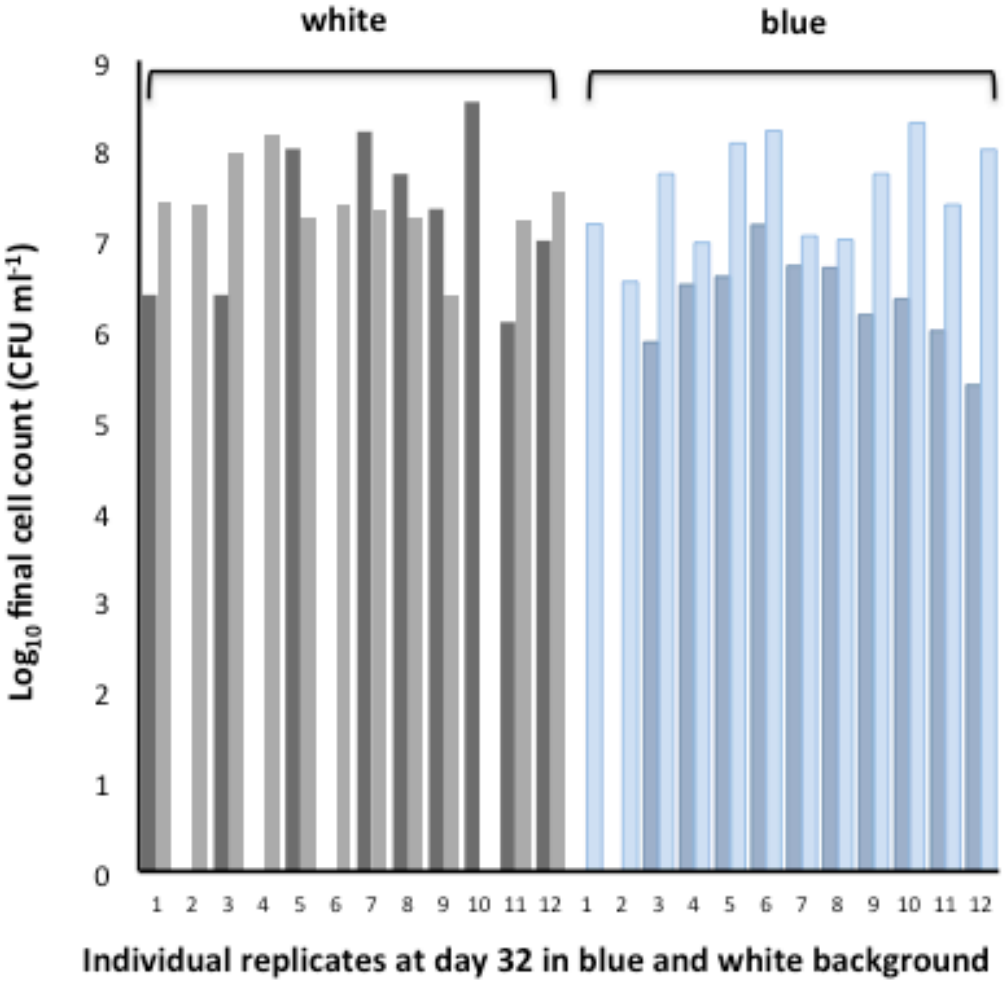

Figure 6.6: The contribution of original plasmid carrying cells and transconjugants to final plasmid carrying cell counts. Original plasmid carriers (■/■) and transconjugants (■/■) (log_{10} CFU ml^{-1}) in the blue and white host backgrounds. Bars represent cell counts at day 32 for individual serial transfer replicates with the conjugative pCT plasmid.

6.3.3 Evolved fitness advantages do not explain plasmid survival

One possible explanation for the increase in proportion of plasmid carriers in serial transfer 2 is some evolutionary change in either the host or the plasmid (or both), which may compensate for the cost of plasmid carriage, or even make plasmid carriage advantageous. In order to examine this, several end-point assays were conducted to give a preliminary comparison of evolved and ancestral lineages and to identify any obvious changes.

One replicate with the non-conjugative plasmid in the blue host background (2b1) had a particularly high final proportion of plasmid carriers, 0.24 (Figure 6.5 A). Competition experiments were conducted to compare the relative fitness of this evolved lineage to the ancestral MG1655 + pCT Δ*trbA* (Figure 6.7). The fitness of plasmid carriers was not significantly different in the evolved 2B1 strain (mean -0.179) than the ancestral strain (mean -0.194; $t = 0.123$, $p = 0.903$).

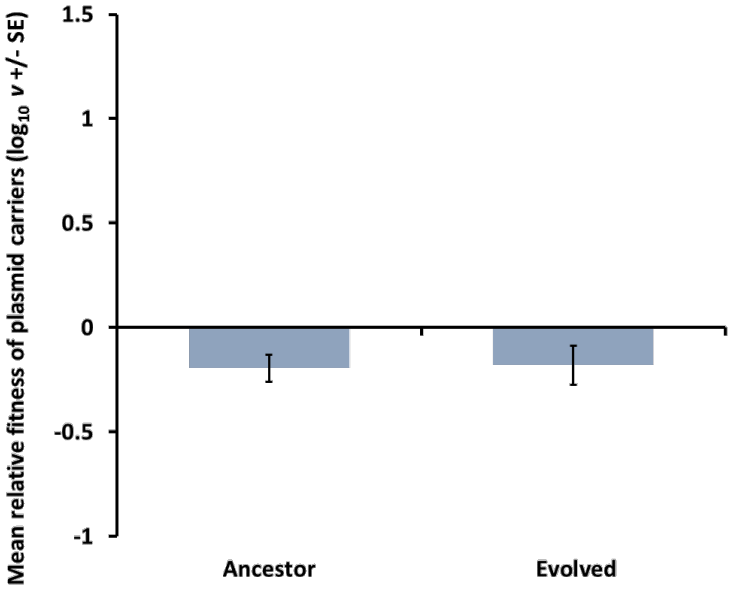

Figure 6.7 The relative fitness of plasmid carriers is not significantly different between ancestral MG1655 + pCT Δ*trbA* (mean -0.194, n = 12) and the evolved derivative 2b1 (mean -0.179, n = 11) in broth competition with plasmid-free MG1655 Δ*lacZYA* ($t = 0.123$, $p = 0.903$).

In addition to the 2b1 evolved strain, competition and mating experiments were conducted with evolved conjugative pCT carrying strains to assess whether any changes in fitness cost or conjugation rate had evolved over the course of the experiment (Figure 6.8 A). Lineages with the highest final proportion of plasmid carriers from each pCT plasmid treatment were selected. For plasmids in the white background, the evolved strains (1a3 and 1b6) did not differ significantly from the ancestor in plasmid carrier fitness ($F = 3.016$, $p = 0.079$). In the blue background, strain 3a6 has significantly higher relative fitness than the ancestor

(t = 3.547, p =2.93 x 10^{-3}), perhaps indicating some plasmid or host evolution in this case. Mating experiments show that white strain 1a3 produced a significantly lower number of transconjugants than its ancestor (t = 2.964, p = 0.010) and that blue strain 3b6 produced a significantly higher number of transconjugants (t = 2.745, p = 0.021) (Figure 6.8 B).

The marked difference between overall numbers of transconjugants in the white and blue background (0.527 and 1.276 transconjugants per donor respectively) is curious (t = 2.075, p = 0.046). The original hypothesis was that plasmids would hitchhike to higher frequency when in the blue background, as the blue strain has slightly higher fitness in broth. The blue strain also appears to be a more vigorous donor which, coupled with higher fitness, could support hitchhiking. However, it is likely that this result is due to observational bias: it is more difficult to detect white transconjugants amongst blue donors than vice versa. In addition, the exclusion of "donor" from the serial transfer 2 model suggests that the significant effect here is unlikely to be genuine (Table 6.3).

The small difference in mean relative fitness between the blue and white ancestral strains should also be noted (-0.39 and 0.68 for the blue and white backgrounds respectively) as it brings into question the result of the fitness assay in Chapter 2, Section 2.3.3. Here, relative fitness was calculated as described by Levin (1988) and is quoted as mean = 1.02 (95% confidence limits: 0.72 – 1.31, n = 6). This is the equivalent of v = 2.08. The large confidence interval and the low number of replicates suggest that the blue strain, MG1655, may not have higher fitness. If this is the case, and there is not a significant fitness difference between the blue and white strains, this could also explain the failure of these experiments to demonstrate hitchhiking: if both strains are of equal fitness neither would be expected to rise disproportionally to dominance.

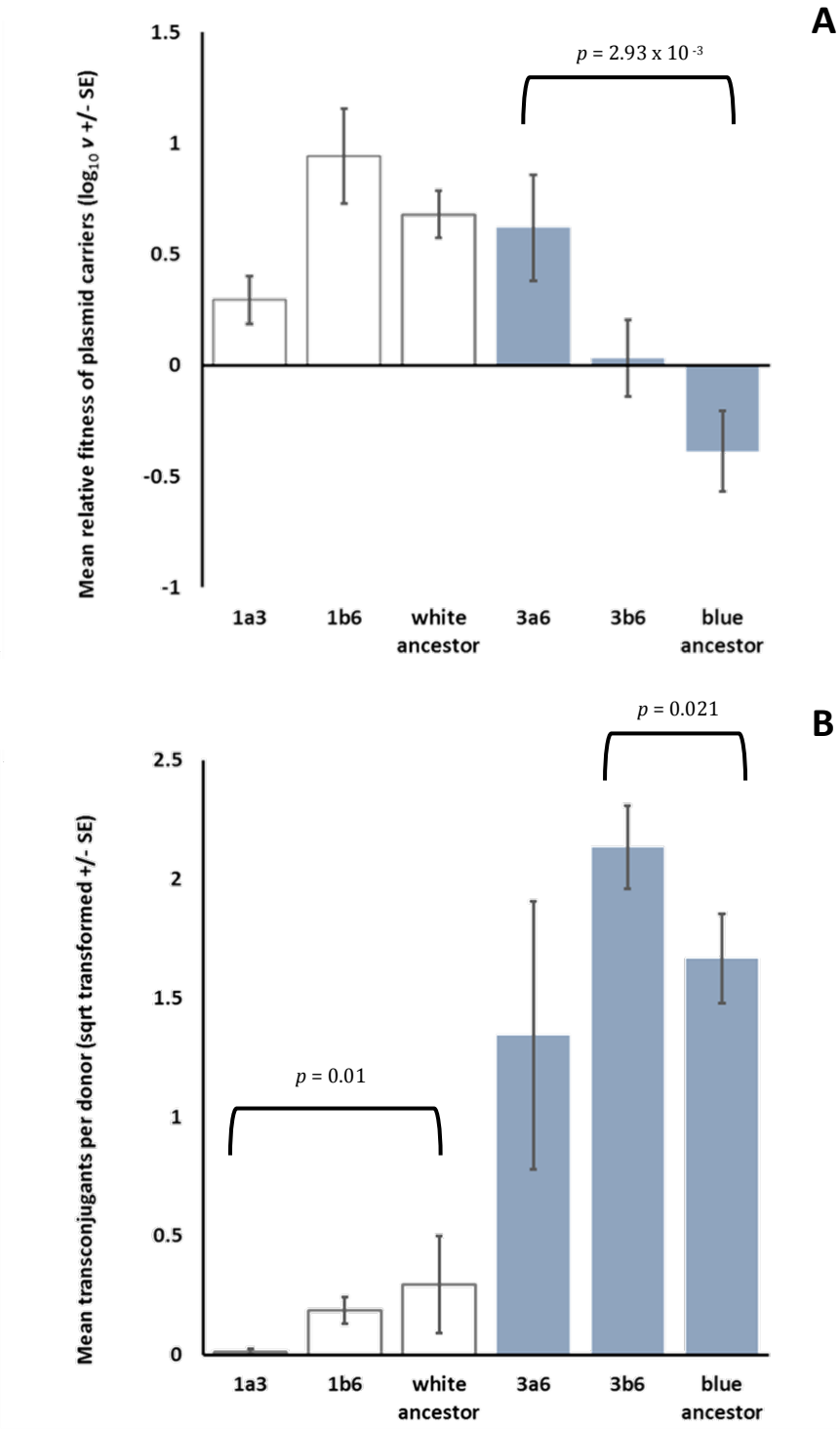

Figure 6.8: Relative fitness and conjugation rate of evolved and ancestral pCT plasmid carrying lineages. The relative fitness **(A)** of plasmid carriers is a significantly different between 3a6 and the blue ancestor. The number of transconjugants per donor **(B)** is significantly lower for 1a3 compared to the white ancestor, significantly higher for 3b6 compared to the blue ancestor.

6.3.4 Plasmid maintenance is frequency dependent

In the first serial transfer experiment (Figure 6.3) the proportion of plasmid carriers decreases from a starting proportion of approximately 0.01 in all cases, with one exception (one replicate using the conjugative plasmid in the low intensity bottleneck). In the subsequent serial transfer (Figure 6.5), the proportion of plasmid carriers increases from very rare (0.00001-0.001) to around 0.01-0.1. Conjugation may explain the maintenance of the plasmid in the pCT mixes by counteracting the effects of segregation and cost of carriage, indeed a large proportion of plasmid carriage is by transconjugants in all but one of the pCT mixes (Figure 6.6). A lack of conjugation could also explain the decline in plasmid carriage in pCT Δ*trbA* mixes between day 19 and day 32. However, conjugation cannot explain the increase in plasmid carriage seen between day 0 and day 19 in the pCT Δ*trbA* mixes, nor can it explain the steeper decline in plasmid carriage in the pCT mixed from day 19. The discrepancy in outcome between the first and second serial transfer experiments indicated that some critical change in plasmid fate was occurring when plasmid carriage reached a proportion between 0.01-0.1, suggesting that plasmid maintenance may have a frequency dependent aspect. In order to assess whether this apparent pattern held true, competitions were set up over a range of initial proportions of plasmid carriage, using the non-conjugative plasmid in both host backgrounds (Figure 6.9). This assay showed no significant effect of host background on relative fitness of plasmid carriers ($F_{(3, 68)} = 2.724$, $p = 0.103$), but in fact a highly significant effect of initial proportion of plasmid carriers ($F_{(1, 70)} = 36.92$, $p = 5.75 \times 10^{-8}$): plasmid carriers do better at low initial proportions. Plasmid carriage has strong negative frequency dependent fitness effects, being beneficial at low frequencies but costly at high frequencies.

It appears that the plasmid- blue background combination is significantly less fit than the plasmid- white background pairing at initial proportions of 0.3 and above (Figure 6.9). Initial analysis suggested no interaction effect between host and initial proportion of plasmid ($F_{(3, 68)} = 1.933$, $p = 0.169$). However, if the initial proportions are grouped into 'low' (0.01 & 0.1) and 'high' (0.3 – 0.9), there is a highly significant interaction effect ($F_{(3, 68)} = 12.348$, $p = 7.9 \times 10^{-4}$) as well as

highly significant individual effects of host background ($F_{(3, 68)} = 8.838$, $p = 4.08 \times 10^{-3}$) and initial proportion plasmid carriers ($F_{(3, 68)} = 270.82$, $p = 2 \times 10^{-16}$). At low frequencies of plasmid carriers the host background is not important: plasmid carriers do better regardless, but at high frequencies there is some epistatic effect between the plasmid and the host: without the plasmid the blue strain is fitter (or no different, see Section 6.3.3 above), but with the plasmid the white strain is fitter. An epistatic interaction like this would completely confound the hitchhiking hypothesis, as the supposedly fitter blue strain would become less fit once it obtained the plasmid.

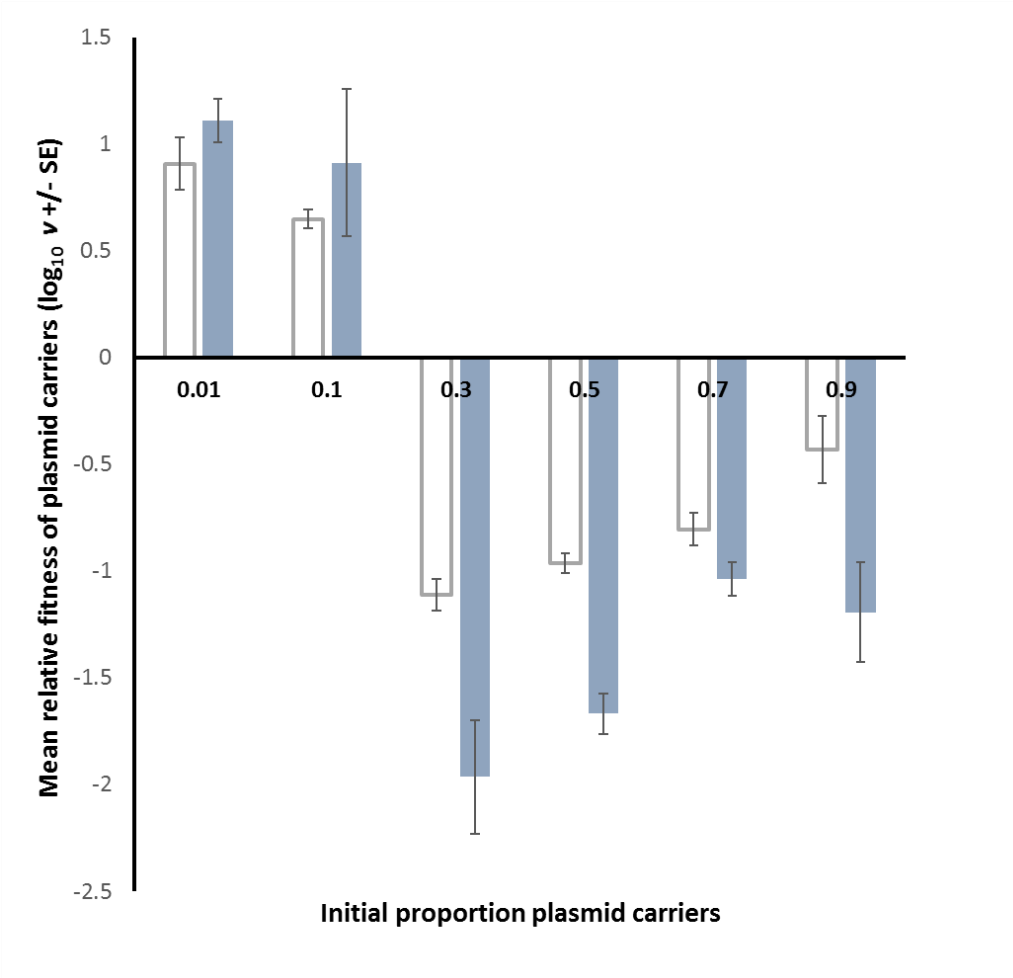

Figure 6.9: Plasmid maintenance is frequency dependent. Mean relative fitness of blue MG1655 (■) or white MG1655 Δ*lacZYA* (□) non-conjugative plasmid carriers, in competition with the opposite plasmid-free strain, over a range of initial proportions. Plasmid carriers show higher relative fitness when the initial proportion is low, but lower fitness at proportions of 0.3 and above.

6.4 Discussion

6.4.1 Bottleneck intensity increases rate of plasmid decline

The importance of population dynamics can be seen clearly in the first serial transfer experiment, with varying bottleneck intensity, and from the effect of the initial proportion of plasmid carriers (see Section 6.4.4 below). The lower the bottleneck intensity, the slower the rate of plasmid decline, regardless of conjugation. Faster plasmid carrier decline will occur with more intense bottlenecks as the number of plasmid carriers surviving to found the next population will be small. Small founding populations in the high intensity bottleneck will lead to more bacterial generations per day (Figure 6.3) and this, coupled with the fitness cost imposed by the plasmid (Chapter 3, Section 3.3.2), will lead to a greater reduction in plasmid frequency in the absence of selection. More intense bottlenecks also increase the impact of genetic drift on the population, meaning that rare plasmid carriers will tend to be lost from the population at higher bottleneck intensities.

It is surprising conjugation has so little impact on plasmid decline: rapid decline of the conjugative plasmid is seen in all bottleneck regimes, although the rate of decline is slightly lower than for the non-conjugative plasmid. It could be argued that the high proportion of plasmid-free recipients at high intensity bottlenecks could facilitate conjugation and somewhat mitigate the effects of population bottlenecking, however, this was not seen here. This result corroborates previous experiments in liquid culture, and suggests that conjugation cannot maintain a plasmid in the face of significant segregational loss but can, to some extent, ameliorate the effect of segregation in comparison to a non-conjugative plasmid (Bahl et al. 2007).

This result highlights the complexity of parasitic plasmid survival: pCT appears to confer no benefit, and is therefore lost from the population. We anticipated that parasitic survival of this large ~90 Kb plasmid would be unlikely, as it appeared to confer no selective advantage (Cottell et al. 2011), and it has had little opportunity for host-plasmid (co-) evolution, as the plasmid was newly

transformed into the host. Although segregation does not appear to play a major role in plasmid decline in this case, the cost of carriage, in the face of competition with a plasmid-free strain and significant population bottlenecks, is too great for conjugation to counteract plasmid decline. The long term coexistence of plasmid-free and plasmid-carrying strains may largely be due to the cost of carriage rather than conjugation or segregational loss (Freter et al. 1983). In addition, mechanistic factors have been shown to affect the contribution of conjugation to plasmid survival, such as the ability to form mating pairs in broth, and the likelihood of donor-recipient meetings at low cell density (Bahl et al. 2007). In this case, low cell numbers after high intensity bottlenecks could inhibit conjugation, which is already limited by the sub-optimal conditions of shaken broth culture.

6.4.2 The importance of conjugation & hitchhiking

One of the key arguments surrounding the parasitic persistence of plasmids is the rate and impact of conjugation: are rates of conjugation high enough to maintain parasitic plasmids? And is this generally true, or are complex, fluctuating environments a prerequisite (Stewart & Levin 1977; Bergstrom et al. 2000; Lili et al. 2007; Ponciano et al. 2007)? Serial transfer 2 shows that pCT can increase from extremely rare, suggesting some benefit to plasmid carriage. Transitory derepression of the conjugative apparatus has been proposed as a mechanism of parasitic plasmid persistence (Lundquist & Levin 1986; Simonsen 1990). Here, the conjugation apparatus of recent transconjugants is temporarily derepressed, facilitating transfer, whilst conjugation is otherwise depressed. The important contribution of transconjugants to plasmid-carrier survival in these experiments could be indicative of this phenomenon, however, the fact that both conjugative and non-conjugative plasmids increased in frequency suggests that there must be other factors in play, and it should also be noted that conjugation rates may be too low for transitory derepression to maintain plasmids (Simonsen 1991; Gordon 1992).

The results of serial transfer 2 indicate that host background does not have a significant effect on plasmid persistence. In contrast to my hypothesis, the fitness of the incoming plasmid-free competitor, relative to the current plasmid host, has no impact on the final proportion of plasmid carrying cells. Plasmids transferring to a fitter strain do not rise disproportionately to success. In fact, the increase in plasmid carriers is seen regardless of host background and plasmid conjugational ability, suggesting that neither conjugation nor hitchhiking are contributing to plasmid persistence (Bergstrom et al. 2000).

6.4.3 The contribution of plasmid and host (co-) evolution

One possible explanation for the increase in plasmid carriers seen in serial transfer 2 is evolution of the plasmid, the host strain or coevolution of the two (Dionisio et al. 2005; Bouma & Lenski 1988; Dahlberg & Chao 2003; Modi & Adams 1991). Evolution may act to facilitate plasmid persistence by altering one or more of the three major parameters affecting plasmid maintenance: conjugation rate, segregation rate and cost of carriage. Preliminary investigations conducted with a selection of evolved strains suggest that, perhaps unsurprisingly, different lineages may have evolved different strategies. However, there is no evidence for strong evolutionary effects in the clones tested. It is of course possible that this is an artifact of the strains selected, and further testing might reveal stronger effects. Alternatively, it could be that comparison of relative fitness is not the most appropriate way to test plasmid and host evolution: re-isolation of evolved plasmids and transformation into ancestral hosts (and vice versa with ancestral plasmids and evolved hosts) would facilitate examination of any host-plasmid coevolution, and assess the contribution of the plasmid and the host to any changes. This would also allow examination of the potential epistatic host - plasmid fitness interactions observed. It is not possible to draw conclusions about the extent or contribution of evolution and/or coevolution to plasmid persistence from this preliminary data. Further phenotypic analyses (as discussed above), and sequencing of plasmids and hosts would allow any changes to be confirmed.

The relative fitness of plasmid carriers in these experiments is not only dependent on the evolution of the plasmid carrying strain, but also on its plasmid free competitor. In order to test for evolutionary changes in the plasmid and host, an experiment in which single strain populations of one host carrying one plasmid, without a competitor, would be more appropriate. The fitness of the evolved strains could be compared to the ancestor by competition experiments using a standard competitor, as described above. In addition, host strains without the plasmid could also be evolved, and their fitness assessed alone and with the addition of an evolved and ancestral plasmid (Table 6.4). This would allow all combinations of evolved and ancestral plasmids and hosts to be combined, allowing evolution of the host strain, the plasmid and coevolution of the two to be identified (see Table 6.4). However, in this experiment the increase in plasmid carrier frequency occurs at the very beginning, suggesting that although evolution may contribute to plasmid maintenance over time, there is some unknown advantage of plasmid carriage from the start of the experiment, which cannot be explained by subsequent adaptation.

Host strain	**Plasmid**
Ancestor	-
Ancestor	Ancestor
Ancestor	Evolved
Evolved (no plasmid)	-
Evolved (cured)	-
Evolved	Evolved
Evolved (no plasmid)	Ancestor
Evolved (no plasmid)	Evolved
Evolved (cured)	Ancestor
Evolved (cured)	Evolved

Table 6.4: Combinations of evolved and ancestral hosts and plasmids to allow evolutionary changes to be attributed to host, plasmid or coevolution. All combinations of host and plasmid would be competed against a standard competitor strain.

6.4.4 Frequency dependent plasmid persistence

Neither conjugation nor coevolution satisfactorily explains the rise in the proportion of plasmid carriers seen in serial transfer 2. I propose two alternative hypotheses: i) a transient fitness benefit for new transconjugants, or ii) a generic benefit of plasmid carriage. Lundquist and Levin (Lundquist & Levin 1986) concede that their data could be explained by a "transitory fitness increase", but suggest that this is biologically unlikely. I propose that this fitness increase could be in the form of increased adaptability to new environment conferred by the plasmid. This would be a transitory benefit, as it would give the plasmid carriers an advantage over their competitors in the first instance, but over time the plasmid free strains would also adapt, and the cost of plasmid carriage may counter the initial benefit of rapid adaptation, leading to the decline we see after day 19. Helling et al. (1981) suggest that beneficial mutations occurring in selectively disadvantaged strains could lead to fluctuating dynamics in competitions, with plasmid carriers gaining a brief advantage, and then being overtaken again by plasmid free competitors. Although in their work the plasmid carriers were in the majority, this could perhaps be compensated for if the plasmid itself were facilitating faster adaptation. While this could explain the decline in plasmid carriers after day 19, as the brief advantage was lost, it does not explain the striking frequency dependent pattern.

Two possible mechanisms could explain the observed frequency dependent fitness of plasmid carriers: i) pCT carries a spiteful trait, which negatively affects the plasmid-free cells; ii) pCT carries some unknown beneficial trait. One example of a negative trait could be a bacteriocin. These are known to be plasmid carried, and in fact the colicins of *E. coli* are exclusively plasmid based (Riley & Wertz 2002). In addition, bacteriocins are known to be involved in spiteful interactions (Gardner et al. 2004) and have a frequency dependent aspect (Chao & Levin 1981). However, Chao & Levin (1981) suggest that colicin-carrying strains only do better when they are common in broth culture, which is the converse of what I have observed with pCT. Similarly, Adams et al (1979), showed frequency dependent patterns with colicin-carrying plasmids, where plasmid carriers do better when their initial frequency is >0.5. Although Cottell

et al. (2011) do not identify any bacteriocin genes in their sequence analysis, I also looked for bacteriocin genes in pCT with the specialized BAGEL3 bacteriocin mining software with no success (de Jong et al. 2006; van Heel et al. 2013). It therefore seems unlikely that this is the cause of the frequency dependent pattern. It could be that there is some undetected novel colicin present, but even if that were true, the patterns of frequency dependence suggest that this is not the cause.

Negative frequency dependence has been observed for beneficial plasmids (Levin 1988). Curiously for this thesis, the beneficial traits carried are often cooperative, public goods-type traits such as mercury resistance or toxin production, and it is these cooperative traits which are subject to frequency dependent selection (Ellis et al. 2007; Raymond et al. 2012). Considering the lack of evidence to indicate a spiteful trait, and the clear support for frequency dependence, it seems that pCT must carry some unknown frequency dependent benefit. This could be in the form of increased adaptability (as suggested above), perhaps through increased mutation rate, or some other (possibly cooperative?) trait.

6.4.5 Conclusions and outlook

This work highlights how little we know about not only plasmid persistence, but also plasmid carried genes and the advantages they may confer (Enne et al. 2004; Bennett & Linton 1986), bringing back into question the existence of cryptic plasmids: as Bergstrom et al. (2000) ask, are they truly cryptic, or are we just yet to discover their advantage? In the case of pCT, these questions might be addressed in future through the production of a knockout library, allowing genes involved in frequency dependence to be identified and studied further.

References

Adams, J. et al., 1979. Frequency-dependent selection for plasmid containing cells of *Escherichia coli*. *Genetics*, (91), pp.627–637.

Bahl, M.I., Hansen, L.H. & Sørensen, S.J., 2007. Impact of conjugal transfer on the stability of IncP-1 plasmid pKJK5 in bacterial populations. *FEMS Microbiology Letters*, 266(2), pp.250–256.

Bennett, P.M. & Linton, A.H., 1986. Do plasmids influence the survival of bacteria? *Journal of Antimicrobial Chemotherapy*, 18, pp.123–126.

Bergstrom, C., Lipsitch, M. & Levin, B., 2000. Natural selection, infectious transfer and the existence conditions for bacterial plasmids. *Genetics*, 155(4), p.1505.

Bouma, J.E. & Lenski, R.E., 1988. Evolution of a bacteria/plasmid association. *Nature*, 335(6188), pp.351–352.

Chao, L. & Levin, B.R., 1981. Structured habitats and the evolution of anticompetitor toxins in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 78(10), pp.6324–6328.

Cottell, J.L. et al., 2011. Complete sequence and molecular epidemiology of IncK epidemic plasmid encoding blaCTX-M-14. *Emerging infectious diseases*, 17(4), pp.645–652.

Dahlberg, C. & Chao, L., 2003. Amelioration of the cost of conjugative plasmid carriage in *Eschericha coli* K12. *Genetics*, 165(4), pp.1641–1649.

de Jong, A. et al., 2006. BAGEL: a web-based bacteriocin genome mining tool. *Nucleic Acids Research*, 34(Web Server), pp.W273–W279.

Dionisio, F. et al., 2005. The evolution of a conjugative plasmid and its ability to increase bacterial fitness. *Biology letters*, 1(2), pp.250–252.

Easter, C.L., Sobecky, P.A. & Helinski, D.R., 1997. Contribution of different segments of the par region to stable maintenance of the broad-host-range plasmid RK2. *Journal of Bacteriology*, 179(20), pp.6472–6479.

Eberhard, W., 1990. Evolution in bacterial plasmids and levels of selection. *The Quarterly review of biology*, 65(1), pp.3–22.

Eberhard, W.G., 2003. Why do bacterial plasmids carry some genes and not others? *Plasmid*, 21(3), pp.167–174.

Ellis, R. et al., 2007. Frequency-dependent advantages of plasmid carriage by Pseudomonas in homogeneous and spatially structured environments. *The ISME journal*, 1(1), pp.92–95.

Enne, V.I. et al., 2004. Enhancement of host fitness by the sul2-coding plasmid p9123 in the absence of selective pressure. *The Journal of antimicrobial chemotherapy*, 53(6), pp.958–963.

Freter, R., Freter, R.R. & Brickner, H., 1983. Experimental and mathematical models of *Escherichia coli* plasmid transfer in vitro and in vivo. *Infection and immunity*, 39(1), pp.60–84.

Galibert, F. et al., 2001. The composite genome of the legume symbiont *Sinorhizobium meliloti*. *Science*, 293(5530), pp.668–672.

Gardner, A., West, S.A. & Buckling, A., 2004. Bacteriocins, spite and virulence. *Proceedings of the Royal Society B: Biological Sciences*, 271(1547), pp.1529–1535.

Gordon, D.M., 1992. Rate of plasmid transfer among *Escherichia coli* strains isolated from natural populations. *Journal of General Microbiology*, 138(1), pp.17–21.

Harrison, E. & Brockhurst, M.A., 2012. Plasmid-mediated horizontal gene transfer is a coevolutionary process. *Trends in Microbiology*, 20(6), pp.262–267.

Helling, R.B., Kinney, T. & Adams, J., 1981. The maintenance of plasmid-containing organisms in populations of *Escherichia coli*. *Journal of General Microbiology*, 123(1), pp.129–141.

Heuer, H., Fox, R.E. & Top, E.M., 2007. Frequent conjugative transfer accelerates adaptation of a broad-host-range plasmid to an unfavorable *Pseudomonas putida* host. *FEMS Microbiology Ecology*, 59(3), pp.738–748.

Levin, B.R., 1988. Frequency-dependent selection in bacterial populations. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 319(1196), pp.459–472.

Levin, B.R. & Stewart, F.M., 1980. The population biology of bacterial plasmids: a priori conditions for the existence of mobilizable nonconjugative factors. *Genetics*, 94(2), pp.425–443.

Lili, L.N., Britton, N.F. & Feil, E.J., 2007. The persistence of parasitic plasmids. *Genetics*, 177, pp.299–405.

Lundquist, P.D. & Levin, B.R., 1986. Transitory derepression and the maintenance of conjugative plasmids. *Genetics*, 113(3), pp.483–497.

Modi, R.I. & Adams, J., 1991. Coevolution in bacterial-plasmid populations. *Evolution*, 45(3), pp.656–667.

Nelder, J.A., 1977. A reformulation of linear models. *Journal of the Royal Statistical Society. Series A (General)*, 140(1), pp.48–77.

Pinheiro, J. et al., 2007. Linear and nonlinear mixed effects models. R package version. Available at: http://CRAN.R-project.org/package=nlme.

Ponciano, J.M. et al., 2007. The population biology of bacterial plasmids: a hidden Markov model approach. *Genetics*, 176(2), pp.957–968.

R Core Team, 2013. R: A language and environment for statistical computing. Available at: http://www.R-project.org.

Raymond, B. et al., 2012. The dynamics of cooperative bacterial virulence in the field. *Science*, 337(6090), pp.85–88.

Riley, M.A. & Wertz, J.E., 2002. Bacteriocins: evolution, ecology, and application. *Annual Review of Microbiology*, 56, pp.117–137.

Ross-Gillespie, A. et al., 2007. Frequency dependence and cooperation: theory and a test with bacteria. *American Naturalist*, 170(3), pp.331–342.

Schlüter, A. et al., 2007. Genomics of IncP-1 antibiotic resistance plasmids isolated from wastewater treatment plants provides evidence for a widely accessible drug resistance gene pool. *FEMS Microbiology Reviews*, 31(4), pp.449–477.

Sia, E.A. et al., 1995. Different relative importances of the par operons and the effect of conjugal transfer on the maintenance of intact promiscuous plasmid RK2. *Journal of Bacteriology*, 177(10), pp.2789–2797.

Simonsen, L., 1990. Dynamics of plasmid transfer on surfaces. *Journal of General Microbiology*, 136(6), pp.1001–1007.

Simonsen, L., 1991. The existence conditions for bacterial plasmids: theory and reality. *Microbial ecology*, 22(1), pp.187–205.

Slater, F.R. et al., 2008. Progress towards understanding the fate of plasmids in bacterial communities. *FEMS Microbiology Ecology*, 66(1), pp.3–13.

Stewart, F.M. & Levin, B.R., 1977. The population biology of bacterial plasmids: *a priori* conditions for the existence of conjugationally transmitted factors. *Genetics*, 87(2), pp.209–228.

Thomas, C.M., 2004. Evolution and population genetics of bacterial plasmids. In B. E. Funnell & G. J. Phillips, eds. *Plasmid Biology*. Washington: American Society for Microbiology (ASM), pp. 509–528.

van Heel, A.J. et al., 2013. BAGEL3: automated identification of genes encoding bacteriocins and (non-)bactericidal posttranslationally modified peptides. *Nucleic Acids Research*, 41(Web Server issue), pp.W448–W453.

Chapter 7: General Conclusion

7.1 Recapitulation of thesis aims

This thesis set out to explore a novel strategy for limiting the spread of antibiotic resistance by exploiting plasmids and bacterial cooperation, and to investigate the evolutionary ecology of plasmids and antibiotic resistance in natural bacterial communities. Four aims were set out in Chapter 1, Section 1.5.2: i) to examine the conditions under which β-lactam resistance is cooperative; ii) to identify and sequence 'cheat' plasmids from natural zoonotic bacterial populations; iii) to investigate the manipulation of cooperative resistance population dynamics using 'cheat' plasmids; iv) to investigate mechanisms of plasmid persistence, in particular the role of conjugation in the maintenance of large naturally occurring plasmids. This concluding chapter brings together the key results stemming from each aim, and highlights the implications of these results for future research.

7.2 Cooperative resistance and cheat plasmids

One major hypothesis was the existence of 'cheat' plasmids; plasmids of the same incompatibility group as a resistance plasmid, which could block the spread of resistance genes by horizontal transfer, and invade resistant populations by social cheating, facilitating the clearance of an infection with antibiotics (see Chapter 1, Section 1.3.2 and Chapter 4 Section 4.1.1). The results of Chapter 3, and the failure to identify 'cheat' plasmids in Chapters 4 or 6, lead me to conclude that 'cheat' plasmids may not be the panacea I had envisaged. With social resistance occurring only under very specific conditions, the presence of susceptible persister cells, there is likely to be limited application for cheat plasmids in a clinical context. Despite this, cross-species antibiotic protection has been seen *in vivo* (Brook et al. 1983; Brook 2004), and the complexity of the host environment and dose variations within tissues may mean that all is not lost for this idea. It is also interesting to note that persistence has been cited as a social trait (Gardner et al. 2007). This opens up the possibility of some mutually

beneficial resistance-persistence interactions, particularly in the context of biofilms where both traits are common.

7.3 Plasmid persistence, frequency-dependence and public goods

Plasmid persistence in competition with plasmid-free strains was explored in Chapter 6, and a strong negative frequency-dependent benefit of pCT carriage was observed, indicating some unknown benefit of plasmid carriage. Beneficial traits carried on plasmids are often cooperative, public goods-type traits, as seen on the plasmids sequenced in Chapter 4: both pCow52/24456 and pCow23 carry siderophore operons, which are commonly associated with both virulence and social traits (West & Buckling 2003), and pCow23 also carries the secreted enterotoxin *espC* (Mellies et al. 2001). Cooperative traits may also be subject to frequency dependent selection, as has been demonstrated for mercury resistance and toxin production (Ellis et al. 2007; Raymond et al. 2012). The fact that public goods genes are more commonly found on plasmids (Nogueira et al. 2009), and that these genes are often subject to frequency dependent selection, suggests that frequency dependent cooperative traits may drive plasmid persistence in some cases: plasmid carrying strains may persist in competition with plasmid free strains if plasmid carriage is beneficial when rare. The unknown cause of the frequency dependent pattern observed in Chapter 6 highlights how little we know about plasmid carried genes and the advantages (or costs) they may confer (Bergstrom et al. 2000).

7.4 Plasmid populations

Naturally occurring plasmids and plasmid populations are still poorly understood as many studies have focussed on clinically or economically relevant plasmids, such as those conferring antibiotic resistance or heavy metal tolerance. The diversity of natural plasmids is highlighted by those sequenced in Chapter 4, which included one completely unique plasmid, pCow63. In addition, the field study in Chapter 5 demonstrated the link between host *E. coli* diversity and plasmid diversity, and found that *E. coli* populations are highly structured. This work found that the *E. coli* population within one pat was likely to contain just

one H-antigen type, and that specific plasmids are associated with specific H-antigens. This clonal host and plasmid population structure within a pat could have implications for cooperative traits (such as the siderophore and toxin genes found on plasmids pCow52/24456 and pCow23), as Hamilton's rule states that cooperation should be directed towards close relatives (Hamilton 1964). Populations carrying the same cooperative plasmid are highly related at the cooperative locus, therefore limiting the potential for exploitation of public goods by cheats (Raymond & Bonsall 2013; Raymond et al. 2012). If plasmid replicon diversity is explained by *E. coli* diversity, and *E. coli* diversity is explained by animal host (pat), then it follows that the plasmid population is highly related within animal hosts, and that there will be limited exploitation of any plasmid carried cooperative traits by cheats. Unfortunately, this high relatedness and increased cooperation may lead to increased virulence, as cooperative genes are also often virulence genes (West & Buckling 2003).

7.5 Conclusions and recommendations

Ten years ago, Thomas (2004) identified three directions for future research into plasmid biology: i) to gain a fuller picture of the plasmid-borne gene pool by focussed sequencing of plasmids from specific species and genera; ii) to understand better how plasmids interact within microbial communities; and iii) to assess the costs and advantages of plasmid carriage that cause plasmids to persist, or not. This thesis has addressed all of these points: firstly, plasmid sequencing (Chapter 4) and replicon typing (Chapter 5) have contributed to our knowledge of the plasmid gene pool of bovine enteric *E. coli*. Secondly, the question of understanding plasmids within microbial communities has been addressed at both a population level, with a non-selective field study of natural plasmids (Chapter 5), and at a smaller-scale, through the investigation of cooperative antibiotic resistance (Chapter 3). And thirdly, the unknown advantages of plasmid carriage have been exemplified by the curious frequency-dependent benefit of pCT carriage observed in the serial transfer experiments (Chapter 6).

To maximise the impact of this work, I propose a number of lines of enquiry: i) the production of a knockout library to identify the genes involved in the frequency dependent benefit of pCT carriage; ii) further sequencing of plasmids isolated in the field study would generate a fuller picture of the plasmid population, and a focus on IncF plasmids would complement the other sequencing work.

Plasmids are major carriers of antibiotic resistance genes, and the spread of ESBLs, in particular CTX-M β-lactamases, has been largely plasmid mediated (Canton et al. 2008). Although the hypothesised 'cheat' plasmids have been elusive, this thesis nonetheless provides insights into the evolutionary ecology of antibiotic resistance plasmids, and also enhances our understanding of plasmid persistence and population structure. I hope that this contribution to knowledge will inform future strategies for the management of plasmid carried antibiotic resistance.

References:

Bergstrom, C., Lipsitch, M. & Levin, B., 2000. Natural selection, infectious transfer and the existence conditions for bacterial plasmids. *Genetics*, 155(4), pp.1505–1519.

Brook, I., 2004. β-lactamase-producing bacteria in mixed infections. *Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, 10(9), pp.777–784.

Brook, I. et al., 1983. In-vivo protection of group A β-haemolytic *Streptococci* from penicillin by β-lactamase-producing *Bacteroides species*. *The Journal of antimicrobial chemotherapy*, 12(6), pp.599–606.

Canton, R. et al., 2008. Prevalence and spread of extended-spectrum β-lactamase-producing *Enterobacteriaceae* in Europe. *Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, 14(s1), pp.144–153.

Ellis, R. et al., 2007. Frequency-dependent advantages of plasmid carriage by *Pseudomonas* in homogeneous and spatially structured environments. *The ISME journal*, 1(1), pp.92–95.

Gardner, A., West, S.A. & Griffin, A.S., 2007. Is bacterial persistence a social trait? *PloS one*, 2(8), p.e752.

Hamilton, W., 1964. The genetical evolution of social behaviour. II. *Journal of theoretical biology*, 7(1), pp.17–52.

Mellies, J.L. et al., 2001. *espC* pathogenicity island of enteropathogenic *Escherichia coli* encodes an enterotoxin. *Infection and immunity*, 69(1), pp.315–324.

Nogueira, T. et al., 2009. Horizontal gene transfer of the secretome drives the evolution of bacterial cooperation and virulence. *Current biology*, 19(20), pp.1683–1691.

Raymond, B. & Bonsall, M.B., 2013. Cooperation and the evolutionary ecology of bacterial virulence: the *Bacillus cereus* group as a novel study system. *BioEssays*, 35(8), pp.706–716.

Raymond, B. et al., 2012. The dynamics of cooperative bacterial virulence in the field. *Science*, 337(6090), pp.85–88.

Thomas, C.M., 2004. Evolution and population genetics of bacterial plasmids. In B. E. Funnell & G. J. Phillips, eds. *Plasmid Biology*. Washington: American Society for Microbiology (ASM), pp. 509–528.

West, S. & Buckling, A., 2003. Cooperation, virulence and siderophore production in bacterial parasites. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1510), pp.37–44.

Appendix 1: RAST annotations of sequenced plasmids

pCow52/24456

Contig	Start	Stop	Strand	Function
1	689	820	+	hypothetical protein
1	2303	1041	-	Putative membrane spanning export protein
1	4495	2348	-	ABC transporter, transmembrane region:ABC transporter:Peptidase C39, bacteriocin processing
1	5676	4492	-	hypothetical protein
1	5976	5818	-	hypothetical protein
1	6664	6792	+	Transposase insE for insertion sequence IS3A/B/C/D/E/fA/fB
1	6960	7397	+	Transposase
1	7436	7837	+	putative integrase
1	8584	8435	-	hypothetical protein
1	8685	9752	+	Glycosyltransferase IroB
1	9892	13551	+	ABC transporter protein IroC
1	13655	14884	+	Trilactone hydrolase IroD
1	14969	15925	+	Periplasmic esterase IroE
1	18147	15970	-	Outer Membrane Siderophore Receptor IroN
1	19952	18918	-	2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I alpha (EC 2.5.1.54)
1	20151	20297	+	hypothetical protein
1	20820	20512	-	ORF2
1	21942	21661	-	stability (stb) locus of IncFII plasmid NR1; similar to SwissProt accession number P11907
1	22170	22048	-	hypothetical protein
1	22202	22357	+	FIG01070796: hypothetical protein
1	22544	22323	-	hypothetical protein
1	23368	23511	+	hypothetical protein
1	24392	23523	-	RepA1
1	24494	24640	+	FIG00642162: hypothetical protein
1	24938	24681	-	Replication regulatory protein repA2 (Protein copB)
1	25408	25178	-	hypothetical protein
1	25768	25421	-	hypothetical protein
1	25795	25941	+	hypothetical protein

1	26582	26061	-	putative nuclease
1	26976	26764	-	hypothetical protein
1	27672	27112	-	IncF plasmid conjugative transfer fertility inhibition protein FinO
1	28473	27727	-	IncF plasmid conjugative transfer pilin acetylase TraX
1	33763	28493	-	IncF plasmid conjugative transfer DNA-nicking and unwinding protein TraI
1	35979	33763	-	IncF plasmid conjugative transfer protein TraD
1	37008	36232	-	IncF plasmid conjugative transfer surface exclusion protein TraT
1	40330	37508	-	IncF plasmid conjugative transfer protein TraG
1	41700	40327	-	IncF plasmid conjugative transfer pilus assembly protein TraH
1	42080	41700	-	IncF plasmid conjugative transfer protein TrbF
1	42348	42067	-	IncF plasmid conjugative transfer protein TrbJ
1	42883	42338	-	IncF plasmid conjugative transfer protein TrbB
1	43151	42870	-	IncF plasmid conjugative transfer protein TraQ
1	43234	43569	+	hypothetical protein
1	43769	43530	-	hypothetical protein
1	43768	43899	+	hypothetical protein
1	44652	43909	-	IncF plasmid conjugative transfer pilus assembly protein TraF
1	44902	44645	-	IncF plasmid conjugative transfer protein TrbE
1	46737	44929	-	IncF plasmid conjugative transfer protein TraN
1	47372	46734	-	IncF plasmid conjugative transfer protein TrbC
1	48373	47381	-	IncF plasmid conjugative transfer pilus assembly protein TraU
1	49002	48370	-	IncF plasmid conjugative transfer pilus assembly protein TraW
1	49385	48999	-	IncF plasmid conjugative transfer protein TrbI
1	52009	49382	-	IncF plasmid conjugative transfer pilus assembly protein TraC
1	52215	52072	-	FIG00641101: hypothetical protein
1	52390	52169	-	IncF plasmid conjugative transfer protein TraR
1	53040	52525	-	IncF plasmid conjugative transfer pilus assembly protein TraV

1	53288	53037	-	IncF plasmid conjugative transfer protein TrbG
1	53497	53300	-	IncF plasmid conjugative transfer protein TrbD
1	54074	53484	-	IncF plasmid conjugative transfer protein TraP
1	55491	54064	-	IncF plasmid conjugative transfer pilus assembly protein TraB
1	56219	55491	-	IncF plasmid conjugative transfer pilus assembly protein TraK
1	56772	56206	-	IncF plasmid conjugative transfer pilus assembly protein TraE
1	57105	56794	-	IncF plasmid conjugative transfer pilus assembly protein TraL
1	57485	57120	-	IncF plasmid conjugative transfer pilin protein TraA
1	57701	57519	-	IncF plasmid conjugative transfer regulator TraY
1	58526	57840	-	IncF plasmid conjugative transfer regulator TraJ
1	59099	58716	-	IncF plasmid conjugative transfer mating signal transduction protein TraM
1	59430	60023	+	X polypeptide
1	60183	60049	-	putative ORF
1	61215	60319	-	unnamed protein product
1	61546	61259	-	hypothetical protein
1	61656	61543	-	Single-stranded DNA-binding protein
1	62146	61946	-	hypothetical protein
1	62276	62112	-	unknown function
1	63521	62802	-	PsiA protein
1	63952	63518	-	PsiB protein
1	64408	64007	-	hypothetical protein
1	65943	64405	-	hypothetical protein
1	66248	66009	-	Putative cytoplasmic protein
1	66751	66299	-	Single-stranded DNA-binding protein
1	66938	66777	-	Single-stranded DNA-binding protein
1	67234	67479	+	hypothetical protein
1	68128	67565	-	Plasmid pO157 DNA, complete sequence
1	69536	68175	-	hypothetical plasmid protein
1	69818	69588	-	YdaB
1	69840	69989	+	hypothetical protein
1	70170	70042	-	Single-stranded DNA-binding protein
1	70235	70363	+	FIG00639891: hypothetical protein
1	70667	70491	-	hypothetical protein
1	71014	70823	-	hypothetical protein
1	71433	71011	-	Orf52 protein
1	71905	71480	-	Putative antirestriction protein
1	72319	72155	-	FIG01049476: hypothetical protein
1	73095	72319	-	FIG00638373: hypothetical protein

1	73575	73141	-	YcgB
1	73810	73589	-	putative cytoplasmic protein
1	74493	73807	-	Adenine-specific methyltransferase (EC 2.1.1.72)
1	74601	74726	+	hypothetical protein
1	74740	74928	+	FIG00639560: hypothetical protein
1	75805	74879	-	Plasmid PO157 DNA, complete sequence
1	77404	76433	-	Chromosome (plasmid) partitioning protein ParB
1	78579	77404	-	Chromosome (plasmid) partitioning protein ParA
1	79913	79158	-	Replication initiation protein RepE
1	80473	80354	-	FIG00638134: hypothetical protein
1	81472	80666	-	Resolvase
1	81778	81473	-	CcdB toxin protein
1	81998	81780	-	CcdA protein (antitoxin to CcdB)
1	82288	82175	-	hypothetical protein
1	82446	82562	+	hypothetical protein
1	85668	82549	-	hypothetical protein
1	85722	86087	+	Transposase insC for insertion element IS2A/D/F/H/I/K
1	86045	86950	+	insertion sequence 2 OrfB protein
1	88160	88591	+	Transposase
1	88659	88805	+	Transposase
1	89758	90735	+	RepFIB replication protein A
1	90930	90811	-	FIG00243898: hypothetical protein
1	91760	91020	-	RepFIB associated resolvase
1	91901	92053	+	hypothetical protein
1	92589	93599	+	hypothetical protein
1	93870	93983	+	hypothetical protein
1	94459	94307	-	hypothetical protein
1	94638	94504	-	Transposase
1	94742	94596	-	hypothetical protein
1	95798	95601	-	hypothetical protein
1	97008	98894	+	membrane; Transport of small molecules: Cations
1	98913	99704	+	hypothetical protein
1	100235	100378	+	hypothetical protein
1	102574	100706	-	outer membrane secretion protein
1	103643	102633	-	hypothetical protein

pCow63

Contig	Start	Stop	Strand	Function
1	1881	1435	-	gpR
1	2821	1976	-	hypothetical protein
1	4262	2832	-	Phage baseplate wedge
1	4633	4259	-	PmgA
1	8421	4639	-	Structural protein P5
1	9257	8436	-	PmgB
1	9331	9450	+	hypothetical protein
1	9646	10401	+	Replication initiation protein RepE
1	11490	10783	-	Tub
1	12057	11506	-	PmgC
1	12603	12112	-	BplB
1	13160	12612	-	PmgG
1	13992	13252	-	hypothetical protein
1	15688	14030	-	hypothetical protein
1	16035	15757	-	hypothetical protein
1	17880	16183	-	hypothetical protein
1	18039	18431	+	hypothetical protein
1	18448	18906	+	hypothetical protein
1	19025	18894	-	hypothetical protein
1	19329	19060	-	hypothetical protein
1	19549	19409	-	hypothetical protein
1	20395	19589	-	UpfB
1	20679	20395	-	UpfC
1	20946	21902	+	Putative stability/partitioning protein encoded within prophage CP-933T
1	21916	22254	+	ParB
1	22542	23183	+	Chromosome partitioning protein ParA
1	23176	23463	+	hypothetical protein
1	23732	25393	+	Glutaminyl-tRNA synthetase (EC 6.1.1.18)
1	26347	25442	-	putative DNA methylase
1	27020	26340	-	hypothetical protein
1	27909	27004	-	DNA recombination-dependent growth factor C
1	28622	27969	-	gp5
1	29599	28619	-	gp6
1	30370	29603	-	hypothetical protein
1	31158	30367	-	gp7
1	31321	31623	+	HigB toxin protein
1	31694	32032	+	HigA protein (antitoxin to HigB)
1	32089	32289	+	Helix-turn-helix motif
1	32311	33039	+	hypothetical protein
1	34266	33073	-	hypothetical protein
1	34588	34259	-	hypothetical protein
1	35570	34917	-	hypothetical protein
1	35839	36462	+	hypothetical protein

1	36483	36875	+	hypothetical protein
1	37183	36872	-	hypothetical protein
1	38169	37243	-	Putative DNA-binding protein Roi of bacteriophage BP-933W
1	38384	38169	-	Icd
1	38865	38698	-	hypothetical protein
1	39775	39197	-	Recombination enhancement function protein
1	39896	40048	+	hypothetical protein
1	40048	40671	+	hypothetical protein
1	40668	40985	+	hypothetical protein
1	40975	41484	+	Eae protein
1	41481	42149	+	putative phage-related protein
1	42149	42472	+	ATP-dependent Clp protease ATP-binding subunit ClpX
1	42469	43254	+	Phage EaE protein
1	43251	43526	+	Phage protein
1	43523	43747	+	hypothetical protein
1	43744	44055	+	hypothetical protein
1	44052	44744	+	Ppp
1	44842	45837	+	Putative cytoplasmic protein
1	45854	46534	+	Phage-related protein
1	46618	46821	+	hypothetical protein
1	46814	47053	+	hypothetical protein
1	47050	47775	+	enterohemolysin 2
1	47777	47971	+	Phage protein
1	47973	48182	+	Phage protein
1	48179	49345	+	hypothetical protein
1	49329	49613	+	hypothetical protein
1	49948	49598	-	hypothetical protein
1	49981	50127	+	DNA polymerase III theta subunit (EC 2.7.7.7)
1	50139	50528	+	Error-prone repair protein UmuD
1	50663	51085	+	hypothetical protein
1	51194	51415	+	Prevent host death protein, Phd antitoxin
1	51415	51795	+	Death on curing protein, Doc toxin
1	51800	51997	+	PdcA
1	52806	52030	-	IsaA
1	53490	52813	-	conserved hypothetical protein protein
1	53999	53505	-	Deoxyuridine 5'-triphosphate nucleotidohydrolase (EC 3.6.1.23)
1	54472	53996	-	hypothetical protein
1	54813	54469	-	hypothetical protein
1	55915	54887	-	GST-loxP-cre recombinase fusion protein
1	56311	57348	+	hypothetical protein
1	58884	57376	-	Terminase B protein
1	60248	58884	-	DNA pacase A subunit
1	60637	60245	-	hypothetical protein
1	61138	60746	-	hypothetical protein

1	61707	61138	-	hypothetical protein
1	62381	62253	-	hypothetical protein
1	63225	62398	-	Putative SPFH domain protein
1	64167	63994	-	hypothetical protein
1	66886	65678	-	hypothetical protein
1	67299	67000	-	hypothetical protein
1	67496	67296	-	hypothetical protein
1	67663	67511	-	hypothetical protein
1	68401	68141	-	hypothetical protein
1	68838	68404	-	Ulx
1	75711	68875	-	N-6 DNA methylase
1	75863	76348	+	FIG00641051: hypothetical protein
1	76578	76381	-	Putative DNA-binding protein Roi of bacteriophage BP-933W
1	77285	76578	-	Putative DNA-binding protein Roi of bacteriophage BP-933W
1	77509	77285	-	Icd
1	78885	77923	-	RepL protein
1	80084	79086	-	Pro
1	81780	80098	-	Prt
1	82197	81865	-	hypothetical protein
1	82799	82242	-	hypothetical protein
1	84289	82925	-	FIG00350536: hypothetical protein
1	84752	84384	-	hypothetical protein
1	86671	84752	-	DarA
1	87284	86664	-	Hdf
1	87720	87274	-	LydB
1	88046	87720	-	LydA
2	101	892	+	Phage tail fiber protein
2	895	1428	+	Phage tail fibers
2	1983	1456	-	hypothetical protein
2	3118	1985	-	Phage tail fibers
3	348	199	-	hypothetical protein
3	576	746	+	hypothetical protein
3	1249	773	-	RepA

pCow23

Contig	Start	Stop	Strand	Function
1	568	29	-	putative transposase
1	1000	689	-	putative transposase
1	1164	1730	+	Resolvase/integrase Bin
1	1938	2468	+	hypothetical protein
1	3289	4056	+	hypothetical protein
1	4318	4193	-	stability (stb) locus of IncFII plasmid NR1
1	4699	4424	-	COG3668: Plasmid stabilization system protein
1	4920	4699	-	hypothetical protein
1	5948	6100	+	hypothetical protein
1	6939	6070	-	RepA1
1	7043	7168	+	hypothetical protein
1	7489	7241	-	Replication regulatory protein repA2 (Protein copB)
1	7669	7791	+	hypothetical protein
1	7922	7773	-	post-segregation killing protein
1	8336	8202	-	hypothetical protein
1	9148	8687	-	putative nuclease
1	9576	9460	-	hypothetical protein
1	9731	9558	-	hypothetical protein
1	10094	9891	-	hypothetical protein
1	10806	10249	-	IncF plasmid conjugative transfer fertility inhibition protein FinO
1	11769	10909	-	Dienelactone hydrolase and related enzymes
1	12574	11828	-	IncF plasmid conjugative transfer pilin acetylase TraX
1	17864	12594	-	IncF plasmid conjugative transfer DNA-nicking and unwinding protein TraI
1	18151	17861	-	IncF plasmid conjugative transfer protein TrbH
1	18173	18571	+	VapC toxin protein
1	20733	18580	-	IncF plasmid conjugative transfer protein TraD
1	20776	20898	+	hypothetical protein
1	21648	20986	-	IncF plasmid conjugative transfer surface exclusion protein TraT
1	21632	21763	+	hypothetical protein
1	22239	21748	-	IncF plasmid conjugative transfer surface exclusion protein TraS
1	25083	22261	-	IncF plasmid conjugative transfer protein TraG
1	26456	25080	-	IncF plasmid conjugative transfer pilus assembly protein TraH
1	26833	26453	-	IncF plasmid conjugative transfer protein TrbF
1	27101	26820	-	IncF plasmid conjugative transfer protein TrbJ
1	27636	27091	-	IncF plasmid conjugative transfer protein TrbB
1	27907	27623	-	IncF plasmid conjugative transfer protein TraQ
1	27988	28302	+	ArtA
1	28651	28304	-	IncF plasmid conjugative transfer protein TrbA

1	29410	28667	-	IncF plasmid conjugative transfer pilus assembly protein TraF
1	29660	29403	-	IncF plasmid conjugative transfer protein TrbE
1	29983	29687	-	IncF plasmid conjugative transfer protein TraN
2	23	163	+	stable plasmid inheritance protein
2	554	165	-	Error-prone, lesion bypass DNA polymerase V (UmuC)
2	1168	812	-	Transposase
2	1598	1371	-	hypothetical protein
2	1748	1918	+	hypothetical protein
2	1935	2888	+	RepFIB replication protein A
2	3083	2964	-	hypothetical protein
2	3913	3173	-	RepFIB associated resolvase
2	4589	5758	+	hypothetical protein
2	5974	6087	+	hypothetical protein
2	7573	7007	-	Transposase
2	8619	7570	-	Transposase
2	9000	8650	-	Insertion Sequence Associated
2	9431	8997	-	hypothetical protein
2	9762	11732	+	Colicin I receptor precursor
2	11739	12530	+	hypothetical protein
2	13111	12593	-	Transposase
2	13106	13222	+	hypothetical protein
2	13246	13458	+	COG2801: Transposase and inactivated derivatives
2	13755	13868	+	hypothetical protein
2	14050	15573	+	reverse transcriptase-like protein
2	15690	16136	+	COG2801: Transposase and inactivated derivatives
2	17495	16671	-	IS2 ORF2
2	17899	17534	-	Transposase insC for insertion element IS2A/D/F/H/I/K
2	18211	17975	-	Per-activated serine protease autotransporter enterotoxin EspC
2	18582	18385	-	Transposase insE for insertion sequence IS3A/B/C/D/E/fA/fB
2	21124	18716	-	putative F1 capsule anchoring protein
2	21861	21142	-	Periplasmic fimbrial chaperone protein
3	1132	278	-	Chromosome (plasmid) partitioning protein ParB
3	2406	1249	-	Chromosome (plasmid) partitioning protein ParA
3	2371	2523	+	hypothetical protein
3	3758	3003	-	Replication initiation protein RepE
3	4338	4219	-	hypothetical protein
3	5338	4532	-	Resolvase
3	5644	5339	-	CcdB toxin protein
3	5864	5646	-	CcdA protein (antitoxin to CcdB)
3	6190	6041	-	hypothetical protein

3	6432	6944	+	putative periplasmic protein
3	7796	6978	-	unknown
3	7836	7988	+	hypothetical protein
3	9051	8278	-	hypothetical protein
3	9564	9064	-	hypothetical protein
3	9753	9640	-	hypothetical protein
3	9829	10059	+	Virulence-associated protein vagC
3	10056	10472	+	VagD
3	13852	10517	-	DNA-binding protein
3	14340	13870	-	hypothetical protein
3	14523	14921	+	Virulence-associated protein vagC
3	14918	15334	+	VagD
3	15409	16974	+	Predicted ATP-dependent endonuclease of the OLD family
3	16959	17981	+	hypothetical protein
3	18201	18037	-	hypothetical protein
4	2009	1371	-	IncF plasmid conjugative transfer protein TrbC
4	3010	2018	-	IncF plasmid conjugative transfer pilus assembly protein TraU
4	3639	3007	-	IncF plasmid conjugative transfer pilus assembly protein TraW
4	4022	3636	-	IncF plasmid conjugative transfer protein TrbI
4	5725	4019	-	IncF plasmid conjugative transfer pilus assembly protein TraC
4	6645	5737	-	IncF plasmid conjugative transfer pilus assembly protein TraC
4	6851	6708	-	hypothetical protein
4	7026	6805	-	IncF plasmid conjugative transfer protein TraR
4	7676	7161	-	IncF plasmid conjugative transfer pilus assembly protein TraV
4	7993	7673	-	IncF plasmid conjugative transfer protein TrbD
4	8567	7980	-	IncF plasmid conjugative transfer protein TraP
4	9984	8557	-	IncF plasmid conjugative transfer pilus assembly protein TraB
4	10712	9984	-	IncF plasmid conjugative transfer pilus assembly protein TraK
4	11265	10699	-	IncF plasmid conjugative transfer pilus assembly protein TraE
4	11598	11287	-	IncF plasmid conjugative transfer pilus assembly protein TraL
4	11965	11603	-	IncF plasmid conjugative transfer pilin protein TraA
4	12393	11998	-	IncF plasmid conjugative transfer regulator TraY
4	13181	12492	-	IncF plasmid conjugative transfer regulator TraJ
4	13499	13368	-	IncF plasmid conjugative transfer mating signal transduction protein TraM

4	13750	13526	-	IncF plasmid conjugative transfer mating signal transduction protein TraM
5	418	933	+	Plasmid PO157 DNA, complete sequence
5	930	1319	+	Plasmid PO157 DNA, complete sequence
5	1458	1270	-	FIG00639560: hypothetical protein
5	1597	1472	-	hypothetical protein
5	1705	2388	+	Adenine-specific methyltransferase (EC 2.1.1.72)
5	2389	2610	+	putative cytoplasmic protein
5	2624	3058	+	YcgB
5	3103	3873	+	FIG00638373: hypothetical protein
5	4159	3989	-	hypothetical protein
5	4290	4715	+	Putative antirestriction protein
5	4762	5184	+	Orf52 protein
5	5181	5372	+	hypothetical protein
5	5859	5578	-	hypothetical protein
5	6057	6185	+	Single-stranded DNA-binding protein
5	6175	6312	+	hypothetical protein
5	6407	6637	+	YdaB
5	6689	8050	+	hypothetical plasmid protein
5	8098	8661	+	hypothetical protein
5	8661	8825	+	hypothetical protein
6	449	174	-	Insertion element protein
6	491	643	+	FIG00642294: hypothetical protein
6	983	2299	+	Transposase
6	2314	3060	+	Transposase
6	3569	7324	+	Per-activated serine protease autotransporter enterotoxin EspC
6	7302	7469	+	Per-activated serine protease autotransporter enterotoxin EspC
7	242	120	-	hypothetical protein
7	2831	738	-	Colicin-D
7	3089	3526	+	entry exclusion protein 1
7	3610	4029	+	entry exclusion protein 2
7	4376	4155	-	hypothetical protein
7	6002	4404	-	hypothetical protein
7	6339	5992	-	mobilization protein MobC
7	6382	6573	+	RNAI modulator protein Rom
7	6997	6884	-	hypothetical protein
8	665	471	-	putative replication regulatory protein
8	918	799	-	hypothetical protein
8	1545	3410	+	Type III restriction-modification system methylation subunit (EC 2.1.1.72)
8	3415	6051	+	Type III restriction-modification system DNA endonuclease res (EC 3.1.21.5)
9	42	689	+	hypothetical protein
9	1449	1069	-	CrcB protein
9	2224	1790	-	Enolase (EC 4.2.1.11)

9	3549	2692	-	Manganese ABC transporter, inner membrane permease protein SitD
9	4403	3546	-	Manganese ABC transporter, inner membrane permease protein SitC
9	5227	4400	-	Manganese ABC transporter, ATP-binding protein SitB
9	6141	5227	-	Manganese ABC transporter, periplasmic-binding protein SitA
10	86	1303	+	Citrate:6-N-acetyl-6-N-hydroxy-L-lysine ligase, alpha subunit (EC 6.3.2.27), aerobactin biosynthesis protein IucA
10	1300	2577	+	L-lysine 6-monooxygenase [NADPH] (EC 1.14.13.59), aerobactin biosynthesis protein IucD
10	2659	4860	+	Aerobactin siderophore receptor IutA
11	77	895	+	IS600 ORF2
11	1234	1362	+	hypothetical protein
11	1363	1662	+	F1845 adhesin operon regulatory protein
11	1670	2212	+	hypothetical protein
11	2227	4671	+	Hypothetical outer membrane usher protein yraJ precursor
12	85	510	+	unknown in ISEc8
12	507	857	+	unknown in ISEc8
12	1244	840	-	hypothetical protein
12	3741	2170	-	Transposase
12	4108	3761	-	orf, conserved hypothetical protein
12	4758	4108	-	Insertion Sequence Associated
12	4854	5009	+	Transposase
13	18	296	+	Putative fimbrial chaperone protein
13	331	822	+	K88 minor fimbrial subunit faeF precursor
13	1060	1896	+	fimbrial antigen FaeG precursor
13	2050	2850	+	K88 minor fimbrial subunit faeH precursor
13	2882	3646	+	K88 minor fimbrial subunit faeI precursor
13	3663	4436	+	K88 minor fimbrial subunit faeJ precursor
14	20	1438	+	putative transposase
14	2258	1482	-	Transposase
14	3283	2261	-	hypothetical protein
14	3383	3967	+	putative transposase
15	1960	2	-	FIG00638906: hypothetical protein
15	2259	2026	-	Putative cytoplasmic protein
15	2861	2322	-	Single-stranded DNA-binding protein
15	3093	2887	-	Single-stranded DNA-binding protein
16	878	48	-	possible AraC-family transcriptional regulator
16	1523	1281	-	putative arylsulfatase regulatory protein
16	1786	1661	-	hypothetical protein
16	1775	2302	+	Transposase
16	2271	2621	+	Transposase
17	129	344	+	Single-stranded DNA-binding protein
17	1493	438	-	invertase

17	1693	1980	+	hypothetical protein
17	2024	2920	+	unnamed protein product
18	46	675	+	Transposase
18	823	701	-	hypothetical protein
18	891	778	-	hypothetical protein
18	1349	1140	-	Dr hemagglutinin AFA-III operon regulatory protein afaF
18	1579	1851	+	IS, phage, Tn; Transposon-related functions
18	1900	2118	+	Insertion element iso-IS1n protein insB
19	823	617	-	FIG00511078: hypothetical protein
20	389	273	-	post-segregation killing protein
20	511	699	+	Protein sok
20	1429	668	-	PsiA protein
20	1860	1426	-	PsiB protein
21	12	446	+	Citrate:6-N-acetyl-6-N-hydroxy-L-lysine ligase, alpha subunit (EC 6.3.2.27), aerobactin biosynthesis protein IucA
21	447	1394	+	N6-hydroxylysine O-acetyltransferase (EC 2.3.1.102), aerobactin biosynthesis protein IucB
23	223	816	+	X polypeptide
23	976	842	-	putative ORF

Appendix 2: Raw plasmid and *E. coli* field data

Total *E. coli* collected per site

Plot	A	CC	D	OC	R	RH	TM	W	WC	SUM
E. coli	140	52	4130	6392	6101	2408	484	488	4288	24483

E. coli H-antigen types by site

	A	CC	D	OC	R	RH	TM	W	WC	SUM
H2	0	0	0	0	4	1	0	2	0	7
H4	0	7	0	40	0	0	0	0	0	47
H5	4	0	1	0	12	0	0	0	0	17
H6	0	0	0	0	0	0	0	0	2	2
H7	1	0	7	0	1	0	0	0	1	10
H8	4	10	4	0	0	0	0	0	0	18
H10	1	0	0	8	19	0	0	0	0	28
H11	0	0	1	1	4	0	0	0	0	6
H12	0	0	1	31	0	19	47	0	20	118
H14	0	0	0	4	0	0	0	0	1	5
H15	0	0	8	0	0	0	0	0	0	8
H16	1	0	2	0	7	0	1	4	0	15
H18	0	0	0	2	0	0	0	0	2	4
H19	1	0	0	3	1	0	0	0	0	5
H20	0	0	0	2	0	0	0	0	0	2
H21	6	0	10	1	1	1	0	0	27	46
H27	0	0	0	0	0	0	1	0	20	21
H28	0	0	4	0	0	30	0	0	0	34
H31	3	0	0	0	1	0	0	7	0	11
H32	0	0	0	0	0	0	0	0	1	1
H33	0	0	0	0	0	9	0	0	0	9
H37	10	0	2	0	1	0	0	1	0	14
H38	7	0	2	0	9	0	0	2	0	20
H39	8	0	0	0	8	0	0	0	0	16
H41	0	0	0	1	0	0	0	0	0	1
H42	0	0	2	0	0	0	0	0	0	2
H48	0	0	0	0	12	0	0	17	0	29
H49	0	1	0	0	14	0	0	5	0	20
H56	1	0	9	1	0	0	0	0	0	11

Plasmids by site

Site	A	CC	D	OC	R	RH	TM	W	WC	SUM
KBBO	0	2	0	5	0	0	0	0	0	7
FIC	0	0	8	0	23	0	0	0	1	32
P	0	14	8	1	0	26	33	0	21	103
FIA	16	0	7	0	60	0	0	21	14	118
FIB	38	17	41	36	79	27	42	23	49	352
Y	1	0	0	1	0	0	0	0	7	9
I1	0	0	10	0	27	0	0	0	0	37
None	8	0	10	53	10	30	1	10	17	139